

6-2016

# On the role of information in educational choice (Doctoral dissertation)

Luca FACCHINELLO

*Singapore Management University*, [lfacchinello@smu.edu.sg](mailto:lfacchinello@smu.edu.sg)

Follow this and additional works at: [https://ink.library.smu.edu.sg/soe\\_research](https://ink.library.smu.edu.sg/soe_research)



Part of the [Education Economics Commons](#)

---

## Citation

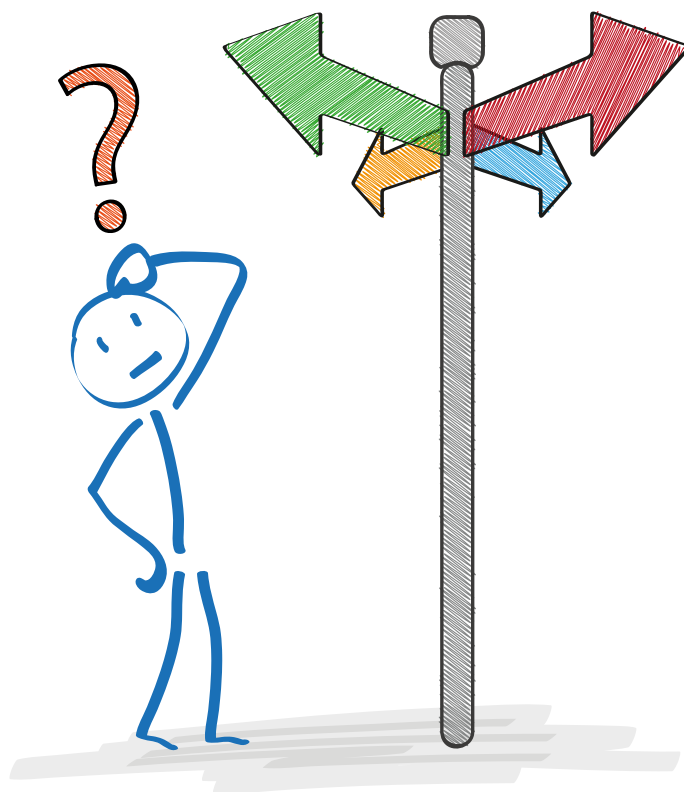
FACCHINELLO, Luca. On the role of information in educational choice (Doctoral dissertation). (2016). 1-200. Research Collection School Of Economics.

**Available at:** [https://ink.library.smu.edu.sg/soe\\_research/2144](https://ink.library.smu.edu.sg/soe_research/2144)

This Book is brought to you for free and open access by the School of Economics at Institutional Knowledge at Singapore Management University. It has been accepted for inclusion in Research Collection School Of Economics by an authorized administrator of Institutional Knowledge at Singapore Management University. For more information, please email [libIR@smu.edu.sg](mailto:libIR@smu.edu.sg).

# On the Role of Information in Educational Choice

Luca Facchinello



## On the Role of Information in Educational Choice

This Ph.D. thesis in Economics consists of three self-contained chapters that investigate how information affects the choices of students.

*The Impact of Early Grading on Academic Choices: Mechanisms and Social Implications* studies theoretically and empirically how the education choices of students with different ability and socioeconomic status respond to early grade assignment.

*Rethinking Education Choices: The Effect of Surveys* investigates whether surveying compulsory school students affects their educational choices, attainment, and long-run labor market outcomes.

*Does Peer Ability Affect Education Choices?* studies whether students assess their academic ability in relation to their peers, and how this relates to their education choices.



### Luca Facchinello

holds a bachelor degree in International Economics from the University of Padua, a master of science in Economics from the University of Bologna, and studied for his Ph.D. at the Stockholm School of Economics. His primary research fields are the Economics of Education and Labor Economics.

# On the Role of Information in Educational Choice

Luca Facchinello

Akademisk avhandling

som för avläggande av ekonomie doktorsexamen  
vid Handelshögskolan i Stockholm  
framläggs för offentlig granskning  
tisdagen den 7 juni 2016, kl 13.15,  
i sal 120, Handelshögskolan, Sveavägen 65, Stockholm



# On the Role of Information in Educational Choice



# On the Role of Information in Educational Choice

Luca Facchinello





Dissertation for the Degree of Doctor of Philosophy, Ph.D.,  
in Economics  
Stockholm School of Economics, 2016

On the Role of Information in Educational Choice

© SSE and Luca Facchinello, 2016

ISBN 978-91-7258-991-9 (printed)

ISBN 978-91-7258-992-6 (pdf)

This book was typeset by the author using L<sup>A</sup>T<sub>E</sub>X.

*Front cover picture:*

© Shutterstock/style-photography, 2016

*Back cover picture:*

Gianna Agostini, 2015

*Printed by:*

Ineko, Göteborg, 2016

*Keywords:*

Grades, academic ability, information, uncertainty, learning, school choice, educational attainment, surveys, peer effects.



*A mia mamma*



# Foreword

This volume is the result of a research project carried out at the Department of Economics at the Stockholm School of Economics (SSE).

This volume is submitted as a doctor's thesis at SSE. In keeping with the policies of SSE, the author has been entirely free to conduct and present his research in the manner of his choosing as an expression of his own ideas.

SSE is grateful for the financial support provided by the Jan Wallander and Tom Hedelius Foundation which has made it possible to fulfill the project.

*Göran Lindqvist*

Director of Research  
Stockholm School of Economics

*Richard Friberg*

Professor and Head of the  
Department of Economics  
Stockholm School of Economics



# Acknowledgements

This thesis is the result of my Ph.D. studies at the Stockholm School of Economics. I am proud of this achievement, and I would like to thank all the people that, in different ways, were involved in the process.

The biggest thanks go to my two advisors, Erik Lindqvist and Juanna Joensen, from whom I learned so much. I really appreciate the encouragement and support they offered me throughout the Ph.D. Thanks to Erik's precious guidance my efforts were always channeled into productive directions. I thank him for his helpful comments, for always being available, and for putting a lot of his time into teaching me how to properly write a research article. Juanna's insightful comments and knowledge of the literature greatly contributed to the dissertation. I thank her for including me in her research project in education, which allowed me to strongly improve the first chapter of the thesis.

I am indebted to Jeff Smith, who sponsored my stay at the department of Economics of the University of Michigan. What I learned from him and the faculty during my fourth year of Ph.D. improved the quality of the thesis.

My stay abroad, and the Ph.D. itself, could have not been possible without the generous financial support of the Jan Wallander and Tom Hedelius Foundation, which I gratefully acknowledge.

I would like to thank Juanna Joensen and Greg Veramendi for their work in the survey paper, and Elena Mattana, Evelina Bonnier, and John Eric Humphries for productive discussions within the education project.

I thank Martina Björkman Nyqvist, Tore Ellingsen, Johanna Wallenius, Federica Romei, Kelly Ragan and Paul Segerstrom for their help in preparing me for the job market. A special thanks goes to Kerem Coşar, for his excellent work throughout the market.

My years in Stockholm would have not been the same without the company of fellow Ph.D. students and friends Arieda Muco, Paola Di Casola, Spyridon

Sichlimiris, Theodoros Rapanos, Ana Maria Ceh, Egle Karmazeine, André Richter, Taneli Mäkinen, Alberto Vesperoni, Abel Schumann, Jakob Almerud, Niels-Jakob Harbo Hansen, Audinga Baltrunaite, Eleonora Freddi, Marta Gigaheddu, Nadiia Lazhevskaya, and Amanda Jakobsson. I really appreciate the “drinks”, parties, academic bickering and pleasant conversations of these years! I also thank Theodoros, Niels-Jakob, Shuhei, Audinga, Mounir and Miri for making the job market experience more pleasant.

While the thesis was written during the Ph.D., I could never have made it here without the support of outstanding teachers along the path. I would like to thank those that never get any glory, but were so important to me. I am thankful to Annamaria Boroso and Tiziana Bernardi for respectively fostering creativity and logical thinking when I was a child, Clotilde Todescan and Mirella Caberlin for providing solid foundations in language and math, and their high school counterparts, Giovanna Viola and Paola Morettin, for keeping up the good work. A special thanks goes to my high school English teacher, Elena Ruffatto, for being such an inspiring teacher and making possible my studies abroad.

My graduate studies in Bologna were fundamental to my advancement to a Ph.D. program. I thank all the LMEC faculty for their exceptional work, and in particular my two advisors, Chiara Monfardini and Andrea Ichino, for guiding my first steps into research.

I could never have made it through the up and downs of the Ph.D. without the support of my family and friends. My mom Antonella and my dad Carmelo were always there for me when I needed it. Our constant TV video-chats made me feel like I never left. I thank my brother Andrea and his wife Laura for keeping me up-to-date with their life and growing family, my niece Emma and my nephew Giovanni for cheering us up with their smiles. A big thank to my friends Jacopo, Marina, Erika and Elisa, who made my stays in San Pietro in Gu feel always too short.

*Stockholm, April 27, 2016*

*Luca Facchinello*

# Contents

<b>Contents</b>	<b>ix</b>
<b>Introduction</b>	<b>1</b>
<b>1 The Impact of Early Grading on Academic Choices: Mechanisms and Social Implications</b>	<b>5</b>
1.1 Introduction . . . . .	6
1.2 Institutional Setup . . . . .	11
1.3 Model . . . . .	17
1.4 Model's Results . . . . .	21
1.5 Empirics . . . . .	33
1.6 Empirical Results . . . . .	37
1.7 Discussion . . . . .	48
1.8 Conclusion . . . . .	49
1.A Numerical Model . . . . .	52
1.B Descriptives . . . . .	77
1.C Refutability Tests . . . . .	96
<b>References</b>	<b>112</b>
<b>2 Rethinking Education Choices: The Effect of Surveys</b>	<b>119</b>
2.1 Introduction . . . . .	120
2.2 Institutional Setting and Empirical Strategy . . . . .	125
2.3 Data . . . . .	133
2.4 Empirical Results . . . . .	138
2.5 Conclusion . . . . .	140
2.A Balancing Tests . . . . .	141
2.B Results . . . . .	154

<b>References</b>	<b>165</b>
<b>3 Does Peer Ability Affect Education Choices?</b>	<b>171</b>
3.1 Introduction . . . . .	172
3.2 Setup . . . . .	174
3.3 Empirics . . . . .	180
3.4 Mechanisms . . . . .	188
3.5 Conclusion . . . . .	192
3.A Indexes . . . . .	193
3.B Rank Deviations . . . . .	197
<b>References</b>	<b>200</b>



# Introduction

When students are uncertain about own ability, information might affect their academic choices. The thesis consists of three self-contained chapters that explore different facets of this theme.

The first chapter studies theoretically and empirically the role of early grade assignment in education choice, focusing in particular on mechanisms. The main argument is that early grading affects differently the choices of students with different academic ability and socioeconomic status.

The second chapter investigates empirically whether repeatedly surveying compulsory school students affects their educational choices, attainment, and long-run labor market outcomes. The basic idea is that educational surveys might contain information relevant for the choices of the students.

The last chapter studies empirically whether, and how, students' choices in compulsory school are affected by peer ability. If students assess their academic ability in relation to their classmates, peer ability might have an effect on their academic choices.

A short summary of each chapter follows.

## The Impact of Early Grading on Academic Choices: Mechanisms and Social Implications

Does early grading affect educational choices? To answer this question, I exploit a curriculum reform which postponed grade assignment in Swedish compulsory schools. The staggered implementation of the reform allows me to identify short- and long-term effects of early grading, for students with different academic ability and socioeconomic status (SES). When graded early on, high-ability students (especially if high-SES) exhibit higher grades in compulsory school, and are more likely to choose academic courses. Low-ability students

react in the opposite way, with particularly negative reactions among low-SES students. High school attainment increases for high-ability low-SES students; college attainment decreases for low-ability low-SES students. None of these effects carry over to the labor market. This suggests that early grades improve the match between early education choices and academic ability, and reduce over-investment in education. I show that the short-term effects are consistent with predictions from a learning model in which children are uncertain about academic ability, have different priors depending on SES, and use grading information to re-optimize educational choices. I find no evidence of demotivating effects for low-ability students, an alternative mechanism through which grades might affect education choices, and the main motivation behind the grading reform.

### Rethinking Education Choices:

#### The Effect of Surveys

(with Juanna Schrøter Joensen and Greg Francisco Veramendi)

Can surveys affect investments in education? This paper examines whether individual education choices and outcomes are affected by a survey posing questions related to expectations and forward-looking behavior. We have administrative data for the whole Swedish population to which an extensive education survey was administered to randomly drawn samples of 3rd graders. This constitutes a randomized social experiment for testing whether responding to survey questions alters behavior. We observe complete educational and labor market histories until the individuals are 31 years old. We have exogenous variation in the timing of first surveys and when an additional survey was administered to parents. The causal effect of the survey on both short- and long-run outcomes is generally not significantly different from zero, independently of parental education. We find, however, that being surveyed increases educational attainment and job stability in the early career for some subpopulations. We will address more specifically heterogeneity of the effect in future research.

### Does Peer Ability Affect Education Choices?

Average classroom ability matters if children assess their ability relative to their peers. I use detailed survey data on a cohort of Swedish 6th graders to estimate

the overall effect of classmates' ability on students' choices in compulsory school. I show that variation in class ability within schools is unrelated to own ability and other determinants of education choice. I find that a one standard deviation increase in average class ability reduces by 2 percentage points the probability of taking an advanced math course in grades 7 to 9. Peer ability does not affect English course choices in grades 7 to 9, and whether students choose academic tracks in high school. I look at underlying mechanisms and show evidence that peer ability negatively affects students' assessment of own ability. The different reduced-form effects on math and English course choices reflect different spillovers in performance: students benefit much more from having high ability peers in English, an interactive subject, than they do in math. Finally peer ability does not seem to affect student's motivation, class interaction and parental support, but positively affects teacher interaction.



# Chapter 1

## The Impact of Early Grading on Academic Choices: Mechanisms and Social Implications<sup>1</sup>

---

<sup>1</sup>A special thanks to my two advisors, Erik Lindqvist and Juanna Joensen, who provided excellent guidance and sound advice throughout the paper. I thank John Bound, Charlie Brown, Kerem Coşar, Susan Dynarski, Tore Ellingsen, Jeffrey Smith and Kevin Stange for their valuable comments. Lastly I thank seminar participants at Stockholm School of Economics, University of Michigan (CIERS and Economics Lunch Seminar), SITE, Brucchi-Luchino and Collegio Carlo Alberto for their useful feedback. I gratefully acknowledge financial support from the Swedish Foundation for Humanities and Social Sciences (Riksbankens Jubileumsfond) grant P12-0968 enabling the data collection for this project. The usual disclaimers apply.

## 1.1 Introduction

While education is traditionally seen in economics as a form of investment with known costs and returns (Becker, 1994; Ben-Porath, 1967), recent models of education choice (e.g., Altonji, 1993) have highlighted the role of uncertainty in educational investment: the expected return of any education choice depends ex-ante on the probability of graduation, and thus on academic ability. When students are uncertain about ability, information, in the form of school grades, might affect their choices.

The role of grades on education choice has been studied almost exclusively at the college level (Stinebrickner & Stinebrickner, 2012; Zafar, 2011). Little is known about how grades affect students at early stages of education, when children have less information on their academic ability, and are still unconstrained by previous choices.

In this paper I investigate how assigning grades in early compulsory school affects educational choices and attainment of Swedish students. To investigate mechanisms I compare the empirical results to the predictions of a sequential choice learning model calibrated to the data.

The institutional setup and the data are particularly suitable to answer the research question. In Sweden, students used to receive the first formal grades in school year 3, at age 10. Grades were based on students' rankings in national standardized tests, and thus provided different information from the test scores students received during the year. In 1969 a reform allowed municipalities to postpone grade assignment to school years 6 or 7. In 1982, a second reform compelled all municipalities to postpone grade assignment to school year 8.<sup>2</sup> The reforms, gradually implemented over time in different municipalities, provide a source of exogenous variation in grade assignment.

I use detailed survey and register data on cohorts born in 1967 and 1972. The 1967 cohort comprises treated students, who were living in municipalities where grading started in middle compulsory school (school year 6), and control students, who lived in municipalities where grades were assigned starting from

---

<sup>2</sup>In 2012 the reform was reversed, and grades in school year 6 were reinstated. Currently grading in school year 4 is being discussed.

late compulsory school (school year 7). Students born in 1972 started receiving grades in late compulsory school (in school year 8) in both treatment and control municipalities. If the education choices of students in treatment and control municipalities trend in the same way over time, it is possible to disentangle the effect of early grade assignment from pre-existing differences between the two sets of municipalities. I provide evidence that trends in educational attainment are the same in treatment and control municipalities, for cohorts who did not receive early grades. I also show that pre-treatment differences in determinants of education appear in general to persist over time.

To guide the empirical analysis, I set up a model of early education choice that captures the most important features of the institutional setup.<sup>3</sup> In the model, ability determines optimal effort and education choices during compulsory school. For low-ability students it is optimal to exert low effort and enroll into vocational high school. For high-ability students it is instead optimal to exert higher effort and enroll into academic education. Children are uncertain about their cognitive ability. Their priors reflect aggregate ability distributions: high-SES children are on average endowed with higher ability than low-SES children.<sup>4</sup> Grades reveal information about true ability, and allow students to re-optimize educational choices. As in the institutional setup, grades can be assigned starting from middle or late compulsory school, while they are never assigned in early compulsory school.

The calibrated model shows that early grade assignment results in better sorting of students into education, that is, in choices closer to first best. However, students with the same ability react differently to the ability signals, due to the different priors about ability. Low (high) SES students who receive low (high) ability signals confirm their priors, and thus react strongly to the information. Students who receive signals inconsistent with their priors form imprecise posteriors, thus their responses are weaker. The model solution implies different reactions to early grading for very low ability students, low ability students, and high ability students. When they receive early grades, students with very low ability increase effort in compulsory school, are more likely to choose vocational high school, and thus less likely to drop out of high school.

---

<sup>3</sup>The model builds on the theoretical framework outlined in Altonji et al (2012)

<sup>4</sup>Régner (2002) discusses biases about ability for low SES students in the psychology literature.

Low-ability students on average reduce effort in compulsory school, and are more likely to choose vocational education paths. These responses appear to be stronger for low-SES students, who are more sensitive to low ability signals. When graded early on, high-ability students increase effort in late compulsory school if they are low-SES, and decrease it if they are high-SES. All high-ability students are more likely to choose academic high school, but only low-SES students increase college attainment as a result of early grading: some high-ability high-SES students fail to access college due to early reductions in effort.

The model guides the empirical analysis: I present the effects of early grades for students with different SES and academic ability. SES is proxied by parental education. Ability is measured using cognitive ability tests administered to both cohorts in school year 6, before grade assignment.

To investigate empirically the effects of early grading on short-term effort, I focus on two outcomes: grades and academic course choices in late compulsory school. Higher grades require higher effort; academic courses are more demanding. Results are broadly consistent with model's predictions: when graded early on, low-ability students, especially if low-SES, receive lower grades and are less likely to choose academic courses in late compulsory school. High-ability students exhibit instead higher grades in late compulsory school, but do not revise course choices. The pattern found in the model is thus reproduced by the data, with the difference that high-ability high-SES students are putting more effort, instead of reducing it.<sup>5</sup>

I consider thereafter effects of early grades on high school choices and attainment. I find an increase in high school enrollment for all students.<sup>6</sup> Contrary to model predictions, I do not observe changes in high school track choice. Why is this the case? I propose as an explanation that preferences for education might attenuate the effects of early grades. My data shows that, controlling for ability, high-SES students' academic high school enrollment rates are 20 percentage points higher than those of low-SES students. At the same time grade differences in late compulsory school between high- and low-SES students

---

<sup>5</sup>The result can be easily reconciled with the model assuming that different college majors require different ability levels.

<sup>6</sup>This is due to the increase in effort during compulsory school for high-ability students. While low-ability students reduced effort, the weakest students could have increased effort when graded early on.



are at most one fourth of a grade: SES appears thus to strongly influence high school choices in Sweden, independently of ability. I find effects on educational attainment only for low-SES students. Early grading leads to a 3 percentage points decrease in college attainment for low-ability low-SES students, and a 6 percentage points increase in high school attainment for high-ability low-SES students (mostly due to a reduction in dropout).

Do the effects found on education carry over to the labor market? I find that early grades do not affect income at ages 33-40, but they increase upward income mobility for low-ability low-SES students. These students showed the strongest reductions in school grades and educational attainment. This suggests that early grades improved the match between early education choices and academic ability, and reduced over-investment in education. Methodologically I confirm the importance of evaluating education policy in the long-run: limiting the analysis to short-term or intermediate education outcomes would have led to different conclusions.

The idea that early grades could motivate/demotivate children in compulsory school was the main motivation behind the grading reform. I empirically investigate this alternative mechanism through which grades could affect education choice. I test for effects of early grades on student motivation and attitudes toward school. These outcomes are measured from survey responses, in the year in which grades were assigned, and in late compulsory school. I find no evidence of early grading discouraging or motivating students, which is consistent with grades simply revealing information to the students.

I conclude that early grades allow students to better sort into education, and thus lead to an increase in efficiency. At the same time the grading policy increases inequality in educational attainment, and reduces effort in compulsory school among low-ability students. The final judgement on the policy depends on the objectives of the policy-maker.

Early grading has relevant effects in the Swedish education system, in which students are explicitly sorted into academic tracks that provide access to college (a *tracked* education system). To what extent do my results generalize to different setups? As knowledge production is cumulative, early education choices constrain late choices for all students (e.g., college preparation affects college enrollment). Assigning grades early on might thus affect students' education

choices and attainment also in non-tracked (*comprehensive*) education systems.<sup>7</sup>

Results are consistent with the learning mechanism outlined by Stinebrickner & Stinebrickner (2012) and Zafar (2011), who find that college students react to grading information. Students who get lower (higher) than expected grades are more (less) likely to drop out/switch to an easier major. A limitation of this literature is that, due to the college setup, it is not possible to tell whether students are learning from grades about academic ability or previous preparation. In my setup grades were assigned when children were 13, so there is less concern that students are learning about previous preparation rather than ability. Moreover I show that students' reactions to grade assignment are consistent with a model in which students learn about ability.

My paper is also related to the grading standards literature, which stresses the role of ability in students' responses to grades. Becker & Rosen (1992) and Betts (1998) show theoretically that higher grading standards encourage high ability students to put more effort, while students below standard might be discouraged. Betts & Grogger (2003) empirically confirm the heterogeneous effects of increasing grading standards at the high school level, while Figlio & Lucas (2004) find that higher standards lead to positive results on test scores, with effects that depend on the ability of the student relative to the class. In my setup untreated students do not observe grades, but only test scores. Absent grades, low-SES students are likely to have lower grading standards than high-SES students (for instance because the difficulty of the tests follow class ability), so that introducing grades should lead to positive effects for high-SES students and negative effects for low-SES students. My results do not confirm this, as I find different reactions to grades within SES.

The grading reform I consider has been previously studied in economics by Sjögren (2010) and in the educational psychology literature by Alli Klapp (2014, 2015).<sup>8</sup> Sjögren's paper uses administrative data to study long-run effects (final education and income) of the overall grading reform. She finds evidence of a positive effect of early grading on educational attainment for girls, and a negative effect for high-SES students. Differences in educational attainment are

---

<sup>7</sup>Early grade assignment has a bigger impact in tracked education systems because students face early choices, and benefit more of timely information about ability. This point has not received much attention in the tracking literature (e.g., Brunello & Checchi, 2007).

<sup>8</sup>Klapp's papers are descriptive regression-control studies

found also before and after the reform took place, which casts some doubts on the robustness of the results. My paper focuses on the mechanisms through which grades affect education choice, and is motivated by a learning model. Results appear to be more robust, as tests for parallel trends in educational attainment do not fail. This is likely due to the different cohorts used: Sjögren needs to assume parallel trends over two decades, while I only need to assume parallel trends within a 5-year period.

The paper proceeds as follows. In Section 1.2 I describe the data, the education system, and the grading reforms. In Section 1.3 I set up the sequential choice learning model that guides the empirical analysis, and illustrate the solution of the model. Section 1.4 discusses the model's results. In Section 1.5 I turn to the empirical analysis, and discuss identification, inference and robustness. Section 1.6 discusses empirical results, while Section 1.7 relates them to the literature. Section 1.8 draws conclusions.

## 1.2 Institutional Setup

### 1.2.1 Data

I use survey data matched to administrative data. The surveys are part of *Evaluation Through Follow-up* (ETF), a longitudinal project which surveys every 5 years representative samples of Swedish students enrolled in compulsory school. I use waves 3 and 4 of the study, corresponding to cohorts born approximately in 1967 and 1972.<sup>9</sup> The 1967 cohort was followed from 1980, when students were in school year 6 (most students were 13 at the time). The 1972 cohort was followed from 1982, when students were in school year 3 (most students were 10 at the time).

Each sample consists of roughly 9000 Swedish compulsory school students (10% of the targeted population) living in 29 (out of 290) municipalities, the lowest administrative division in Sweden. Whole classes were systematically sampled from municipalities, and the same municipalities were extracted in both waves.<sup>10</sup> The final sample is thus a repeated cross-section, which allows me to implement a difference in differences identification strategy.

---

<sup>9</sup>In the following I will refer to the two samples as 1967 and 1972 cohort.

<sup>10</sup>Municipalities are drawn using stratified sampling. Strata are defined by population, fraction of left-wing voters, fraction working in the public sector and fraction of immigrants.

The survey data contains relevant information for the analysis. First, sampled students took standard intelligence tests in verbal, logical and spatial ability in school year 6, before end-of-the-year grades were assigned. The tests are exactly the same for both cohorts, which grants comparability of the intelligence measures over time. At the time of the tests students were 13, a point in which IQ should have already stabilized (Cunha & Heckman, 2009). I can thus investigate the effects of early grading using proper measures of ability, rather than previous performance measures. Second, grades and course choices in compulsory school are recored from school registers. This allows me to inspect the effect of early grading right after grades were assigned. Third, children filled in detailed surveys in school years 6 and 10 (the first year of high school). They were asked questions about own ability, courses and high school track choices, well being and motivation in school. I use children responses about stress, anxiety, and motivation as outcomes to understand whether early grades had motivating/demotivating effects on the children, a main concern in the policy debate. Finally, parents were surveyed when children received their first survey. They were asked questions about school choices and priorities. This evidence helps to understand whether and to what extent choices of parents living in early grading municipalities differ from those of parents living in municipalities where early grades were abolished.

I match to the sample high quality register data from Statistics Sweden. For both cohorts I observe parental education, income and demographics. These variables allow me to test for compositional change in the sample, and allow to increase precision in the main specification. The registers record educational attainment, income, and income mobility at ages 33-40 for both cohorts. This allows me to evaluate how the short- and medium-run effects of early grading transmit to the labor market.

### 1.2.2 The Education System

Table 1.1 summarizes the Swedish educational setup for the two cohorts in my sample. Compulsory school (*Grundskola*) started at age 7 and lasted 9 years. It was formally divided in three stages, that could also entail physically

---

The three biggest municipalities in Sweden (Stockholm, Malmö, Gothenburg) are always part of the sample. Further details on the sampling scheme can be found in Emanuelsson (1979).

changing schools: early compulsory school (grades 1-3), middle compulsory school (grades 4-6), and late compulsory school (grades 7-9). Standardized end-of-the-year grades were released at the end of each education cycle, and in every year during late compulsory school. Early grades were over time abolished. The next section provides details about the grades and the grading reforms.

Table 1.1: Structure of Swedish education

	Compulsory school		Non Compulsory school	
	Early and Middle	Late	High School	College
Age	7-12	13-15	16-19	<b>Selection:</b> - HS track - GPA or SweSAT
School Year	1-6	7-9	10-12	
Grades	(3), (6)	(7), 8, 9	10-12	
Choices	-	General or advanced courses	Vocational or academic track	<b>Funding:</b> - Free tuition and grants - Loans for living expenses
Selection	-	-	GPA and course choices	

The education system was tracked. In the spring of school year 6 children had to choose whether to take math and English at the advanced or general level in the next school year. Academic electives provided better preparation for academic tracks in high school. Students were allowed to switch course type over time. At the end of compulsory school, students could enroll in either academic or vocational high school tracks. Vocational tracks lasted two years, provided professional training, and did not allow direct access to college. Academic high school lasted three or four years, prepared for college, and was selective.<sup>11</sup>

After academic high school graduation (or taking one more year of high school after vocational school) students became eligible to apply to college.

<sup>11</sup>A high grade 9 GPA and advanced math electives in compulsory school could be used as admission requirement.

A student quota, set by the government, limited access to college. Slots were competitively assigned to the students with highest GPA or *SweSAT* (a college entry test similar to the American SAT).<sup>12</sup> College was tuition-free, and a mix of grants and income-contingent loans allowed admitted students to pay for living expenses. Higher education was thus both meritocratic and competitive. Appendix 1.B.2 presents detailed evidence on education choices and attainment for sampled students.

### 1.2.3 Grades and the Reform

Standardized grades in math, English and Swedish were assigned at the end of specific school years during compulsory school. Grades were norm-referenced at the national level: they represented student performance with reference to the whole student cohort.<sup>13</sup> Given that only homework and test scores were assigned during the school year, grades provided students with additional information about school performance. In particular grades provided a first idea of their chances of admission to college, which was restricted by a quota system.

The school year in which grades were first assigned was over time postponed from school year 3, when students were 10, to school year 8, when they were 15.<sup>14</sup> Up to 1968 grades were assigned in school years 3, 6, 7, 8 and 9. In 1969 a curriculum reform (*Curriculum Lgr 69*) allowed municipal school boards to abolish “early” grades, that is, grades in school years 3 and 6. As a substitute for the abolished grades the reform introduced parent-teacher conferences, non-compulsory biannual meetings in which teachers evaluated pupil improvement over the year. Sjögren (2010) reports that supporters of early grade abolition were concerned about early grades harming low SES or poorly performing

---

<sup>12</sup>Öckert (2002) reports that around 50% of the students were rejected admission to college in the period I study, confirming the selective nature of Swedish higher education.

<sup>13</sup>Tests were corrected by the teachers. The government used the scores to determine the national grade distribution. When assigning final grades, teachers could deviate from test scores, if they thought the student test performance did not reflect proficiency.

<sup>14</sup>In 2012 grades were reintroduced in school year 6, and the government is considering assigning grades also in school year 4.

students. The idea behind the reform was that of making the class environment less competitive and more inclusive.

Since 1969 more and more municipalities took the chance to abolish grades in the early school years, but the issue was contentious. Left parties (Social Democrats and Communists) in general favored early grades abolition, while right-wing parties (Center party and Moderate Party) leaned towards keeping the early grades (this is confirmed in Figure 1.B.2 on page 86). In the end the government, led by a socialist majority, chose to abolish “early grading” in all municipalities: starting from 1982 (*Curriculum Lgr 80*) end-of-the-year grades were released only starting from school year 8, when children were 15.

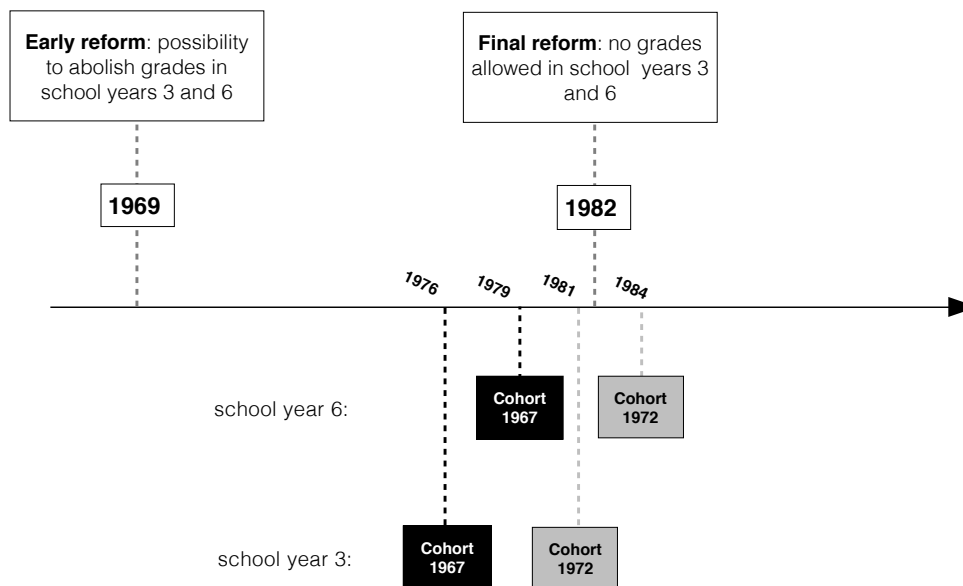


Figure 1.1: Grading reform timeline and sampled cohorts

Figure 1.1 shows in a timeline how the reforms affected the two cohorts in the sample. Half of the municipalities in the 1967 cohort sample were assigning grades in school year 6, while the rest had abolished them.<sup>15</sup> Grade assignment in school year 3 is not recorded in my data for this cohort, as the children were followed from school year 6. Using information provided in Sjögren (2010), I can assume that the municipalities assigning grades in year 6 could have been also assigning grades in year 3. However municipalities not assigning grades in year 6 should have also abolished grades in school year 3. No municipality

<sup>15</sup>Figure 1.B.3 shows in a map which sampled municipalities were assigning early grades.

in the 1972 cohort sample was assigning grades in school year 6. While the final reform was effective the year after the children born in this cohort were in school year 3, my data reports that no grades were assigned in school year 3. Finally end-of-the-year grades were assigned for all cohorts and municipalities in school years 8 and 9.<sup>16</sup>

Table 1.2: Grade assignment

	1967 Cohort	1972 Cohort
<b>Early Grading Municipalities (Treatment)</b>	(Year 3) Year 6	
	Year 7	
	Year 8	Year 8
	Year 9	Year 9
<b>Late Grading Municipalities (Control)</b>	Year 7	
	Year 8	Year 8
	Year 9	Year 9

In the following I emphasize the role of grades in school year 6, rather than school year 3. First, treatment status in my analysis is based on grade assignment in school year 6. Second, grades at the age of 13 are arguably more relevant than grades at age 10, the end of early compulsory school. At that stage, grades might be more informative of effort, or preferences for education, rather than academic ability. Finally, after school year 6 students had to choose whether to take math and English at the general or advanced level. Grades in school year 6 should thus be more relevant for education choices.

Table 1.2 summarizes the grading structure. I label “treatment municipalities” those municipalities that were assigning grades in school year 6 before the final reform, “control municipalities” those not assigning grades in school year 6 before the final reform. The treatment is thus receiving early grades in school year 6 (and potentially 3), which holds for students born in 1967 who lived in early grading municipalities.

<sup>16</sup>Differently from earlier school years, they were assigned two times per year, at the end of each semester. Details are taken from Skolverket.



### 1.3 Model

The model presented in this section investigates how early grading affects students' education choices and attainment when grades convey information about ability. The qualitative predictions of the model are compared to empirical results in Section 1.6.

#### 1.3.1 Structure of the Model

The model focuses on the link between early education choices, educational attainment, and lifetime income. I model explicitly early phases of education, and treat non-compulsory education and the labor market as realizations. The structure of the model is illustrated in Figure 1.2.

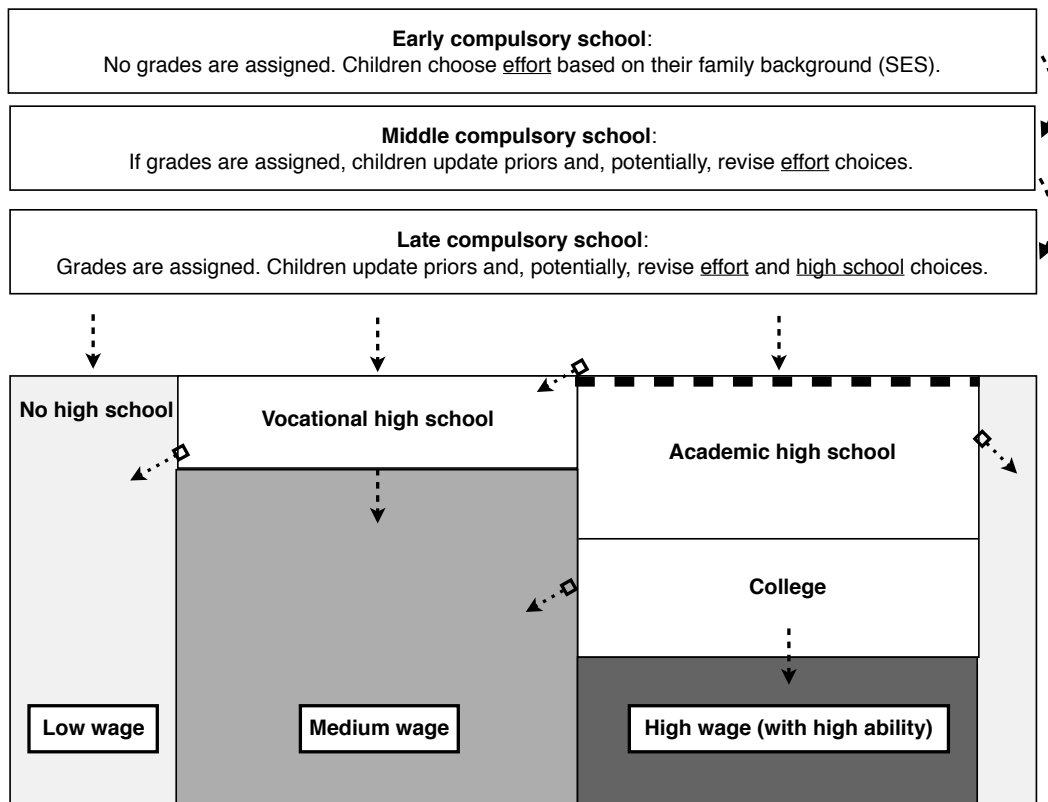


Figure 1.2: Structure of the model

Compulsory education is divided, as in my setup, into three periods: early compulsory school ( $t_1$ ), middle compulsory school ( $t_2$ ), and late compulsory school ( $t_3$ ). In each period student  $i$  chooses how much effort to exert:  $e_{it} \in \{1, 2, 3\}$ . Effort choices and academic ability ( $a_i$ ) determine the stock of knowledge ( $k_{it}$ ) the student accumulates:<sup>17</sup>

$$k_{it} = \omega_t(\alpha a_i + \beta e_{it}) + \delta k_{it-1}. \quad (1.1)$$

After the end of compulsory education students have three choices. They can go to work ( $E_1$ ) and earn low wages ( $w_1$ ). They can enroll into vocational high school ( $E_2$ ) and study for two years, or enroll into academic high school ( $E_3$ ) and study for 3 years. Both high school tracks grant medium wage ( $w_2 = w_3$ ) upon graduation. Academic high school is the only option that gives access to college ( $E_4$ ), which lasts 4 years and grants upon graduation wages that increase with academic ability:  $w_4 = f(a_i)$ .<sup>18</sup>

Completing higher levels of education and accessing academic high school requires higher knowledge (and thus higher ability and effort) at the end of period 3. Notice that knowledge is here not productive per se, as wages fully depend on attained education and ability. This is consistent with a signaling model where employers are uncertain about workers' ability, but observe attained education (Spence, 1974). Higher education is attained in equilibrium only by high-ability workers, who fetch higher wages in the market. The knowledge thresholds at time 3 are the following:

$$\bar{k}^{E_2} < \underline{k}^{E_3} < \bar{k}^{E_3} < \bar{k}^{E_4}, \quad (1.2)$$

where  $\underline{k}^{E_j}$  and  $\bar{k}^{E_j}$  are respectively the entry and attainment requirements for education level  $E_j$ . Failure to meet the thresholds results in dropout (assumed at the midpoint of each education level), and thus in foregone earnings. Given that high school grants the same wage independently of track, it is optimal to enter academic high school only under the expectation to be able to complete college.<sup>19</sup> Academic ability indeed determines optimal education and effort

<sup>17</sup>The three stages of education have different lengths in my setup. Weights  $\omega_t$  adjust the length of each stage to mimic the actual setup.

<sup>18</sup>As this is a stylized model, returns to education do not reflect the substantial wage heterogeneity documented in the literature (Arcidiacono, 2004; Hussey et al, 2011).

<sup>19</sup>Notice also that I do not model entry to college, and simply consider people staying out as college dropouts.

choices. Low-ability students ( $a_i \in \{1, 2, 3\}$ ) optimally choose vocational school, and put during compulsory school levels of effort inversely proportional to their ability: to reach the same education level, a weaker student needs to exert higher effort in school. The optimal education choice of high-ability students ( $a_i \in \{4, 5\}$ ) is academic high school, and thus college. To attain college education they need to exert higher effort in compulsory school.

Students are uncertain about academic ability:  $\tilde{a}_{it} \sim f_t(a_i)$ . They have priors reflecting the ability distribution by SES:  $f_1(a_i) = f(a_i|SES)$ . In particular low-SES students have on average lower ability than high-SES students. Grades are unbiased ability signals, and allow students to update their priors about academic ability. They can be assigned in middle compulsory school, and are always assigned in late compulsory school, before students choose high school track:

$$g_{i2} = d(a_i + \epsilon_2) \text{ with } \epsilon_2 \sim N(0, \sigma_2^2(\epsilon)), \quad (1.3)$$

$$g_{i3} = d(a_i + \epsilon_3) \text{ with } \epsilon_3 \sim N(0, \sigma_3^2(\epsilon)), \quad (1.4)$$

where  $d$  is a function that maps the normal values into the discrete ability scale. Grades assigned in late compulsory school are more precise than grades assigned in middle compulsory school:  $\sigma_2^2(\epsilon) > \sigma_3^2(\epsilon)$ . This reflects the fact that more grades are assigned in the last period of compulsory school.

Table 1.3: Information structure

	Early grades	Late grades
$f_1(a_i)$	$f(a_i SES)$	$f(a_i SES)$
$f_2(a_i)$	$f(a_i g_{i2}, SES)$	$f(a_i SES)$
$f_3(a_i)$	$f(a_i g_{i3}, g_{i2}, SES)$	$f(a_i g_{i3}, SES)$

Table 1.3 makes explicit the information structure in the three periods. If early grades are not assigned in period 2, students' beliefs remain unchanged:  $f_2(a_i) = f(a_i|SES)$ . Otherwise they are updated:  $f_2(a_i) = f(a_i|g_{i2}, SES)$ . In period 3 grades are always assigned, so that  $f_3(a_i) = f(a_i|g_{i3}, SES)$  if no grades are assigned in middle compulsory school, and  $f_3(a_i) = f(a_i|g_{i3}, g_{i2}, SES)$  with early grades. Finally, when students update their priors about ability in period

$\tau$ , they revise their beliefs about accumulated or future knowledge at time  $t$ :

$$\tilde{k}_{it,\tau} = \sum_{j=1}^5 P_{\tau}(a_i = j) \times \left[ \omega_t(\alpha j + \beta e_{it}) + \delta \tilde{k}_{it-1,\tau} \right]. \quad (1.5)$$

### 1.3.2 Optimal Choice

Given their information about ability in period  $\tau$ , students consider optimal effort and education choices in any subsequent period  $t \geq \tau$ . They choose the education level that gives the highest utility, and the associated optimal effort level  $e_{it,\tau}^{E\star}$ :

$$e_{it,\tau}^{E\star} = \arg \max_{e_{it,\tau}^{E\star}} \{ V_i^{E_1}(e_{it}^{E_1\star}), V_{i,\tau}^{E_2}(e_{it,\tau}^{E_2\star}), V_{i,\tau}^{E_3}(e_{it,\tau}^{E_3\star}) \}. \quad (1.6)$$

The value of compulsory school,  $V_i^1$ , does not depend on ability, and is thus not indexed by time. It is maximized when effort is set to the lowest level, so it is a constant:  $V_i^{1\star} = k$ . As vocational and academic high school have access and attainment requirements, values  $V_{i,\tau}^{E_2}$  and  $V_{i,\tau}^{E_3}$  depend on students' beliefs about ability. They are indexed by the time index  $\tau$ , as  $\tilde{k}_{i3,\tau}$  changes when new information is revealed:

$$V_{i,\tau}^{E_2} = \sum_{t=\tau}^3 C(e_{it,\tau}^{E_2\star}) + P(\tilde{k}_{i3,\tau} \geq \bar{k}^{E_2}) \times U((L-2) \times w_2) + P(\tilde{k}_{i3,\tau} < \bar{k}^{E_2}) U((L-1) \times w_1) \quad (1.7)$$

$$\begin{aligned} V_{i,\tau}^{E_3} = & \sum_{t=\tau}^3 C(e_{it,\tau}^{E_3\star}) + P(\tilde{k}_{i3,\tau} \geq \bar{k}^{E_4}) \times U((L-7) \times w_3(a_i)) + \\ & P(\bar{k}^{E_3} \leq \tilde{k}_{i3,\tau} < \bar{k}^{E_4}) \times U((L-5) \times w_2) \\ & + P(\underline{k}^{E_3} \leq \tilde{k}_{i3,\tau} < \bar{k}^{E_3}) \times U((L-2) \times w_1) + \\ & P(\tilde{k}_{i3,\tau} < \underline{k}^{E_3}) \times \\ & [P(\tilde{k}_{i3,\tau} \geq \bar{k}^{E_2}) \times U((L-2) \times w_2) + P(\tilde{k}_{i3,\tau} < \bar{k}^{E_2}) \times U((L-1) \times w_1)] . \end{aligned} \quad (1.8)$$

$C$  is a convex cost function,  $U$  is a concave utility function, and  $L$  is the number of working years.

### The effect of grades

When students are assigned grades they update their priors in the direction of their true ability level. Figures 1.A.5 to 1.A.9 in Appendix 1.A.2 show priors

and posterior distributions of ability after grades are assigned. Updating can have two effects: an “*income*” and a “*substitution*” effect. When the student realizes she has higher (lower) ability than expected, she revises the level of knowledge accumulated upward (downward). Provided the optimal education choice has not changed, the student will need to put weakly less (more) effort to reach the level of non-compulsory education she was targeting, an “*income effect*”:

$$\left. \frac{\partial e_{it}^{E_s^*}}{\partial \tilde{a}_{it}} \right|_{E_t^* = E_{t-1}^*} = \frac{\partial \tilde{k}_{i3}^t}{\partial \tilde{a}_{it}} \times \frac{\partial e_{it}^{E_s^*}}{\partial \tilde{k}_{i3}^t} \leq 0. \quad (1.9)$$

If after observing the signal expected ability is high (low) enough to alter optimal educational choice, the student will instead revise effort choices upward (downward), a “*substitution effect*”:<sup>20</sup>

$$\left. \frac{\partial e_{it}^{E_s^*}}{\partial \tilde{a}_{it}} \right|_{E_t^* \neq E_{t-1}^*} = \frac{\partial \tilde{k}_{i3}^t}{\partial \tilde{a}_{it}} \times \frac{\partial e_{it}^{E_s^*}}{\partial \tilde{k}_{i3}^t} \geq 0. \quad (1.10)$$

## 1.4 Model’s Results

Before discussing the model’s predictions, it is important to be clear about the purpose of the model. First, the model is meant to qualitatively assess the effect of early grades in the specific setup I consider. I calibrate to the data the key parameters of the model, ability distributions and education payoffs. I set thresholds for educational attainment such that higher levels of education require both higher ability and effort. Parameters with no direct counterpart in the data (knowledge production function, precision of grade signals, and value function parameters) are fixed to specific values.<sup>21</sup> Appendix 1.A.1 contains further details on calibration, and provides evidence on model assumptions. Second, I do not estimate the model. While this might be an interesting direction for future research, my aim here is to generate qualitative predictions of the effect of early grading in a learning model, rather than fitting the data.

---

<sup>20</sup>Higher education levels always require higher knowledge.

<sup>21</sup>Results remain qualitatively the same when slightly changing the parameters. Extreme parameterizations lead to different predictions, but are also inconsistent with the data observed.

Table 1.4: Optimal choices under full information

	$a_i$	$e_{i1}$	$e_{i2}$	$e_{i3}$	$E$	$V_{i,1}^E$
Low-ability	1	Medium	Medium	Medium	Vocational	106.53
	2	Medium	Medium	Low	Vocational	112.63
	3	Low	High	Low	Vocational	115.84
High-ability	4	High	High	Medium	Academic	126.49
	5	Medium	High	Medium	Academic	155.48

I solve numerically the model under three different information setups: late grade assignment, early grade assignment and, as a benchmark, full information.<sup>22</sup> In Table 1.4 I show as a reference optimal effort and education choices by ability level under full information. The “*income effect*” is clear for both low- and high-ability students: for higher levels of ability it is optimal to put less effort. The “*substitution effect*” appears when ability changes from 3 to 4: students need to put higher effort early on in order to be able to attain college education.

#### 1.4.1 Effort in Compulsory School

Figure 1.3 shows optimal effort choices in  $t_1$ , before grades are assigned. The fact that additional information will arrive in  $t_2$  might change effort choices before grades are released. This is not the case in the model. Under uncertainty about ability, it is always optimal for both low- and high-SES students to keep effort at a medium level. This is due to three reasons. First, uncertainty favors higher effort early on: putting low effort in the beginning might actually prevent the student from entering academic high school, and thus college. Second, even if the student learns that she is high-ability in time, she would then need to compensate for previous low effort levels: effort cost is convex, so this behavior would not be optimal. Finally, knowledge production is cumulative, so it is better to exert higher effort early on, when effort is more productive.

In middle compulsory school students can be assigned grades. Figure 1.A.10 in Appendix 1.A.3 compares posterior distributions of ability for low and high SES students who get the same grades in  $t_2$ . While all students update priors in the right directions, updates differ by SES. Low (high) SES students who receive

<sup>22</sup>Appendix 1.A.2 presents the simulation and solution methods.

low (high) grades confirm their priors, and thus their posterior distributions have higher densities on low (high) ability levels. Students who receive grades different from their priors form instead posterior distributions with higher weight on intermediate values of ability. These posteriors are thus also less precise.

Figure 1.4 shows the effect of early grading on effort choices in  $t_2$ , by aggregate (low or high) ability and SES. Results for each ability level, reported in Appendix 1.A.3, are useful to better interpret the aggregate picture, so in the following I refer to both pictures. Early grading changes optimal behavior in middle compulsory school only for high-SES students: students who observe signals consistent with high-ability put higher effort (see Figure 1.A.11). As shown in table 1.4, this is consistent with optimal education choice: for high-ability students it is optimal to exert high effort in middle compulsory school, and then reduce effort in late compulsory school. Low-SES students do not react differently at this stage, independently of ability. Their priors are set lower, and hence posteriors about ability are less sensitive to the high grades they observe.

In  $t_3$  all students are graded. Figure 1.5 shows that high-SES students with high-ability strongly react to the additional grades, and put lower effort. Together with the reaction in middle school, this can be overall interpreted as a negative “*income effect*”. High-ability low-SES students react to early grades in the opposite way: they increase effort. Against their priors, these students realize they are high-ability. They thus switch education and effort choices (“*a substitution effect*”). Low-ability students reduce effort when graded early on. Figure 1.A.12 shows that the strongest reductions are found among low-SES students. The aggregate effect for low-ability students masks a positive “*income effect*” among lowest ability students. Figure 1.A.12 shows that, when graded early on, these students put more effort to reach the same education level they targeted (an “*income effect*”). This effect is strongest among low-SES students.

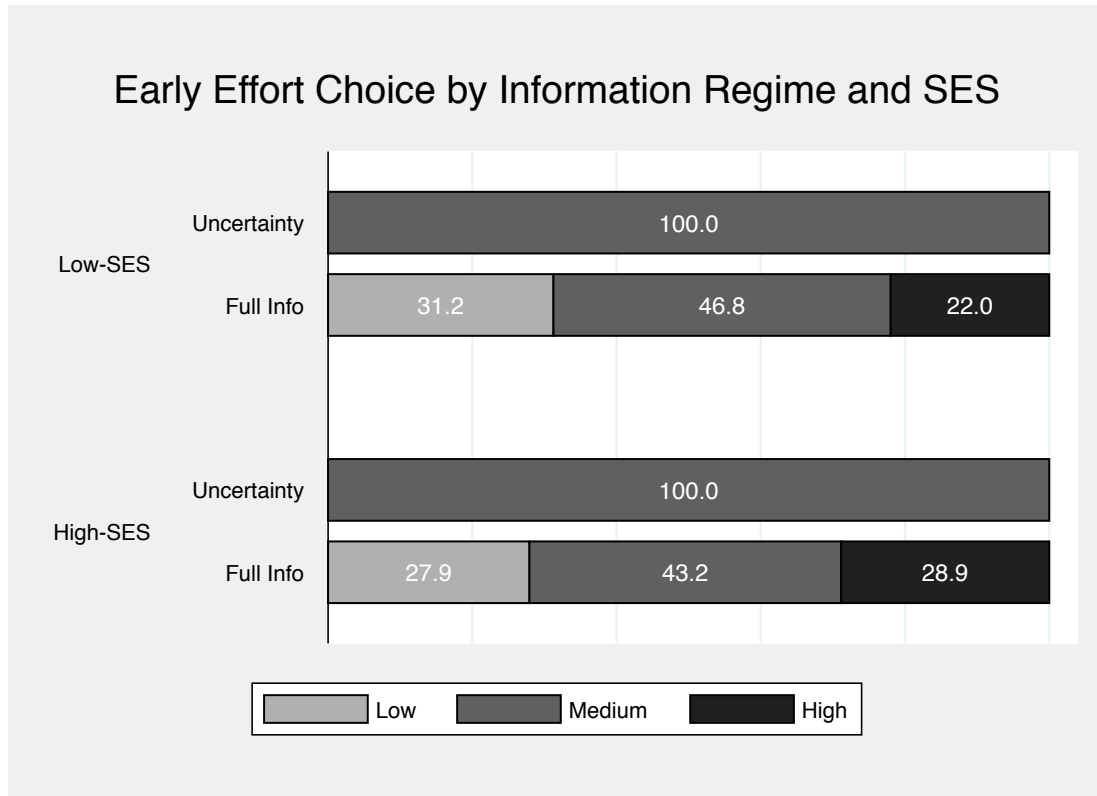


Figure 1.3: Early effort choice by SES and information regime

Note: The Figure plots effort distributions in early compulsory school. Since in this period no grades are assigned, choices are the same for both low and high ability students, and can only differ by SES. Assigning early or late grades does not change effort choices in  $t_1$ .



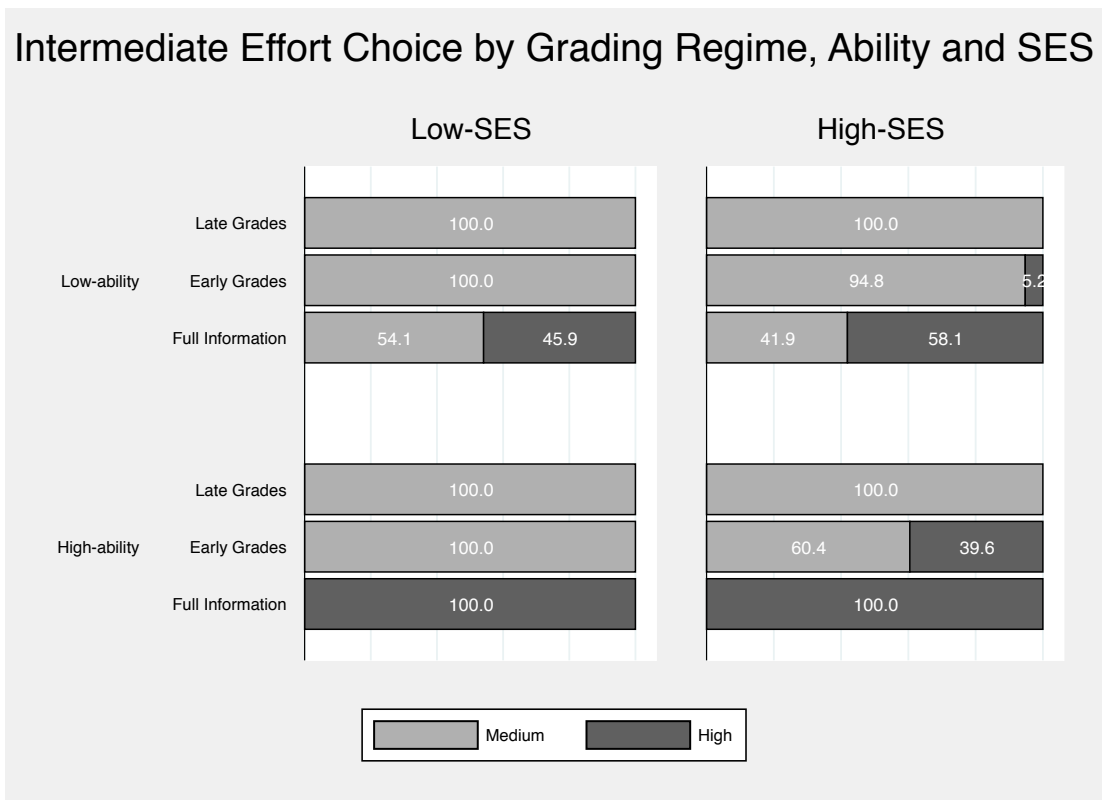


Figure 1.4: Intermediate effort choice by aggregate ability and SES for different grading regimes

Note: The Figure plots effort distributions in middle compulsory school. Results are presented for high-ability students, for whom it is optimal to follow an academic education path, and low-ability students, whose optimal choice is vocational high school. SES affects students's priors about ability, and thus optimal choices.

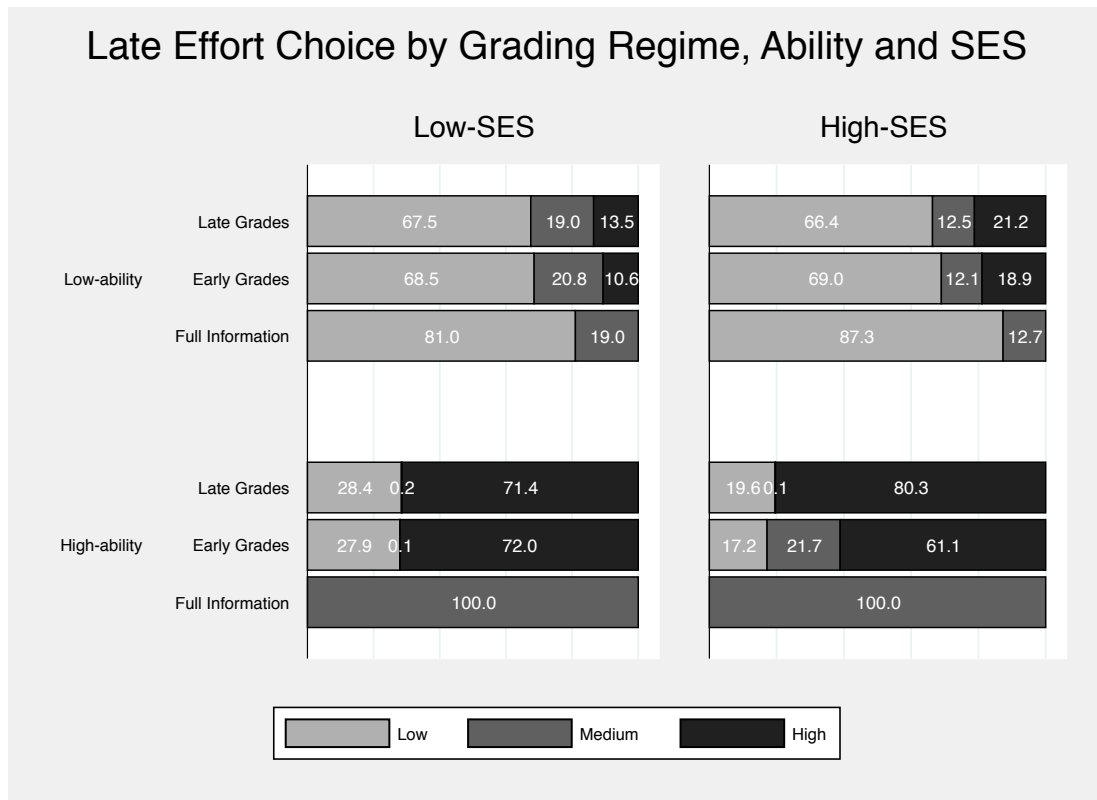


Figure 1.5: Late effort choice by aggregate ability and SES for different grading regimes

Note: The Figure plots effort distributions in late compulsory school. Results are presented for high-ability students, for whom it is optimal to follow an academic education path, and low-ability students, whose optimal choice is vocational high school. SES affects students's priors about ability, and thus optimal choices.

### 1.4.2 Education

Figures 1.6 and 1.A.13 show high school choices in the three information regimes. When graded early on, low-ability students are less likely to choose academic education paths. The effect is stronger for low-SES students. All high-ability students are instead more likely to choose academic high school with early grades. Among students with high (but not top) ability, the reaction is stronger for high-SES students.

Figures 1.7 and 1.A.14 show final education distributions for the different grading setups. The effects of early grades mirror those observed for education choice. The main difference is that some high (but not top) ability students with high-SES fail to attain college, and only complete academic high school. These students observed signals consistent with top ability early on, lowered effort, and thus failed to graduate from college (see Figure 1.A.16). No such effect is found for low-SES students, who are actually less likely to dropout of both high school (see Figure 1.A.15) and college (see Figure 1.A.16).

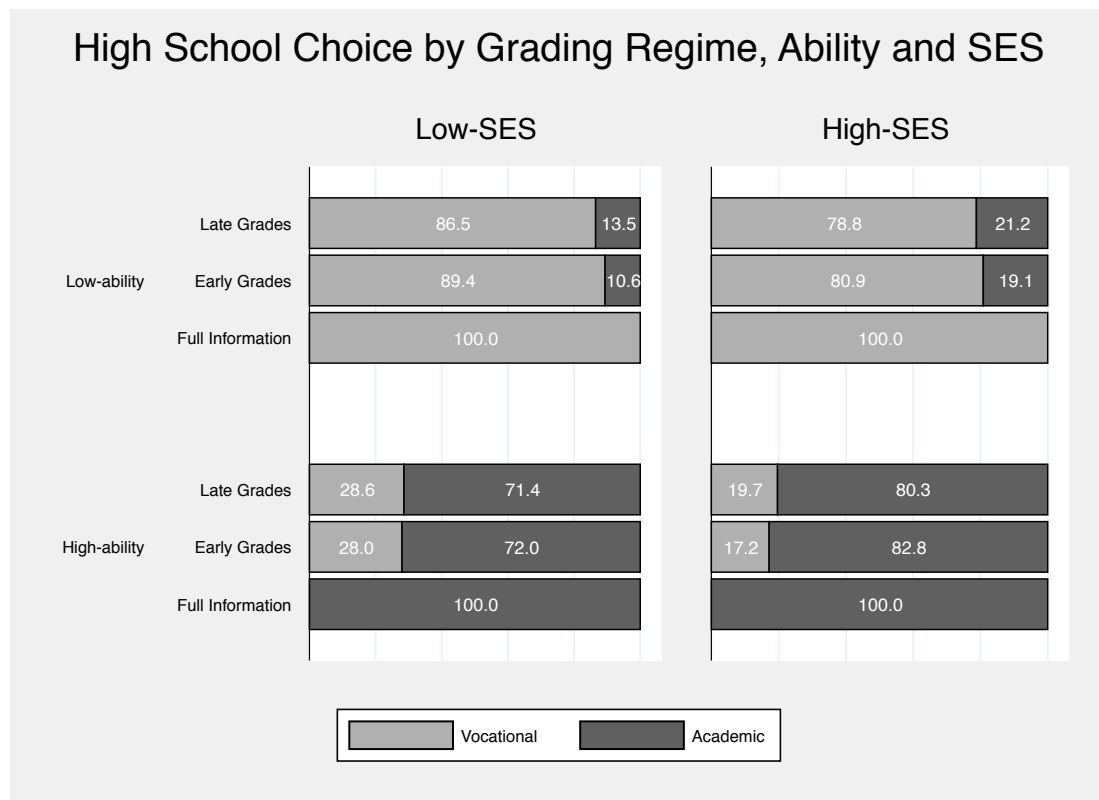


Figure 1.6: High school choice by aggregate ability and SES for different grading regimes

Note: The Figure plots high school choice distributions. Results are presented for high-ability students, for whom it is optimal to follow an academic education path, and low-ability students, whose optimal choice is vocational high school. SES affects students's priors about ability, and thus optimal choices.

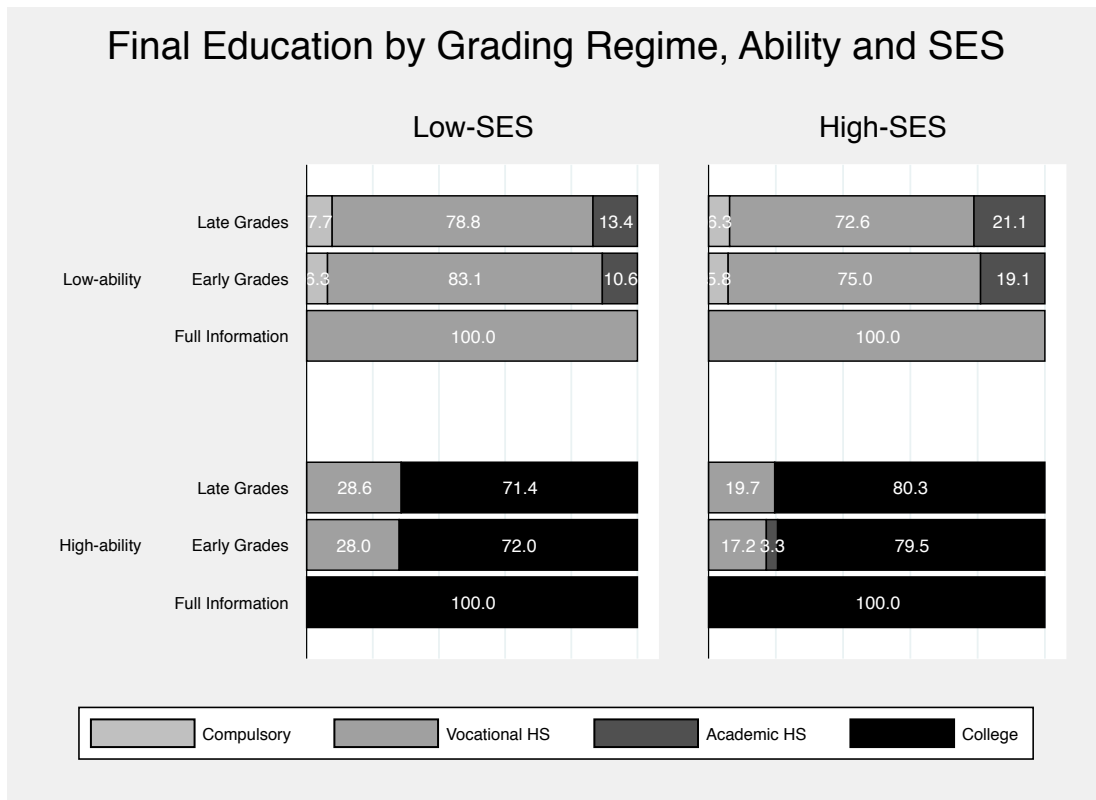


Figure 1.7: Final education by aggregate ability and SES for different grading regimes

Note: The Figure plots final education distributions. Results are presented for high-ability students, for whom it is optimal to follow an academic education path, and low-ability students, whose optimal choice is vocational high school. SES affects students's priors about ability, and thus optimal choices.

### 1.4.3 Summary of Results

In Table 1.5 I summarize the effects of early grading (the treatment) on education choices, educational attainment, and income. The effects are reported for each ability and SES group, and are compared to the baseline scenario (late grading, in brackets).

In general early grades lead to an overall reduction in effort.<sup>23</sup> Only high-ability low-SES students - for whom positive “*substitution effects*” prevail - increase effort when graded early on. While the mean reduction in effort is the same for all low-ability students, effects are qualitatively different by SES. As seen before, there are weaker negative “*income effects*” for high-SES students, and stronger positive and negative “*income effects*” for low-SES students. The biggest negative “*income effect*” on effort is found for high-ability high-SES students, who are the most sensitive to high grade signals.

Staying out of high school is never optimal in the model, even when students realize they have lower ability than expected. The rational choice for these low-ability students is to enroll into vocational school, and later dropout if they fall short of the required preparation. This does not change with early grades, which instead have nontrivial effects on high school track choices: all low ability-students are less likely to enroll into academic tracks, and the opposite is true for high-ability students. Positive reactions are strongest for high-SES students, while negative reactions are more pronounced for low-SES students.

Even if they reduced effort after observing early grades, all low-ability students benefit of the early information, due to the different choices taken at the end of compulsory school: dropout rates decrease, in particular for low-SES students. This translates one to one into an increase in high school attainment. Finally, college attainment increases for high-ability low-SES students, and slightly decreases for high-ability high-SES students.

In the long-run the effects of early grades on education translate into small increases in lifetime income for all students, with the exception of high-ability high-SES students. Effects on income are pretty small, and close to 0. Most of the gains in utility are due to the early reductions in effort, so that early grading improves on average the welfare of all students.

All in all the simulations show that assigning grades earlier leads to choices

---

<sup>23</sup>I take a weighted average of middle and late effort choices in order to provide a more complete picture on the effects of early grades on effort choice.

and education outcomes more consistent with academic ability, with responses differing by SES. Lowest ability students are more likely to increase effort when graded early on, especially if low-SES. Low to medium ability students reduce effort in compulsory school, in particular if low-SES, but are more likely to choose vocational tracks, which they are able to complete. High-ability low-SES students increase effort in compulsory school, are more likely to choose academic paths, and to attain college. For high-ability high-SES students, “*income effects*” tend to prevail: these students put less effort when they observe high grades, which leads some of them to fail to graduate from college.

Table 1.5: Summary of the effects of early grade assignment

Outcome:	Low-Ability			High-Ability	
	All Sample	Low-SES	High-SES	Low-SES	High-SES
<i>Effort in late compulsory school</i>	-0.03 [1.92]	-0.03 [1.55]	-0.03 [1.62]	0.01 [2.36]	-0.07 [2.51]
<i>HS Enrollment</i>	0.00 [1.00]	0.00 [1.00]	0.00 [1.00]	0.00 [1.00]	0.00 [1.00]
<i>Academic track HS Enrollment</i>	-0.01 [0.40]	-0.03 [0.13]	-0.02 [0.21]	0.01 [0.71]	0.02 [0.80]
<i>HS Dropout</i>	-0.01 [0.04]	-0.01 [0.08]	-0.00 [0.06]	0.00 [0.00]	0.00 [0.00]
<i>Attains HS</i>	0.01 [0.96]	0.01 [0.92]	0.00 [0.94]	0.00 [1.00]	0.00 [1.00]
<i>Attains College</i>	-0.00 [0.30]	0.00 [0.00]	0.00 [0.00]	0.01 [0.71]	-0.01 [0.80]
<i>Income (0-1 scale)</i>	0.00 [0.75]	0.00 [0.67]	0.00 [0.68]	0.00 [0.85]	-0.00 [0.88]
<i>Utility</i>	0.51 [117.01]	0.62 [107.83]	0.39 [107.20]	0.23 [128.91]	0.67 [133.17]

Values in brackets represent outcomes when only late grades are assigned. Effort is defined on a 1-3 scale (1 is low effort). Income is a measure of lifetime income, and assumes everybody starts working right after finishing their education or dropping out.



## 1.5 Empirics

In this section I discuss identification of the effect of early grading on education choices. I then briefly discuss inference in my setup, and lastly provide evidence on identifying assumptions.

### 1.5.1 Identification

The decision to assign early grades was taken by municipal school boards, and, as previously discussed, correlates with the political color of the municipality. Treatment assignment is thus likely not random with respect to education outcomes. A simple comparison of outcomes between grading and non-grading municipalities would pick up systematic differences between the two sets of municipalities, and thus bias OLS.

In Appendix 1.B.3 I test for differences in pre-treatment variables between graded and non-graded municipalities in the 1967 cohort. Table 1.B.7 shows that in graded municipalities children are less likely to be foreign born, score better in the ability tests, and are less likely to switch classes over compulsory school. In terms of school level variables (changes of teachers, class size, kindergarten) there are no big differences, in line with the homogeneous nature of Swedish education. Parents in grading municipalities (Tables 1.B.8 to 1.B.12) are less likely to divorce and more likely to be married. They are slightly poorer, less educated, and more likely to be employed in low-skill jobs or agriculture. When asked about how they chose math and English courses, and the priorities of Swedish education, parents give very similar answers. The only differences, the weight they put on the role of parents in school choice and critical thinking in school, do not seem to imply a different preference for children educational attainment. Altogether it appears that there are some small differences in determinants of education choice between the two sets of municipalities. The differences in parental education seem to reflect a different structure of the economy, rather than different preferences for education.

A simple cross-sectional comparison of outcomes for treated and untreated municipalities would likely lead to a negative bias, due to the pre-existing differences between treated and control units. Given that I observe treatment and control group before and after the final reform, when early grades were abolished, I can “control” for any persistent difference between the two sets of municipalities. If outcomes trend in the same way in the two municipalities

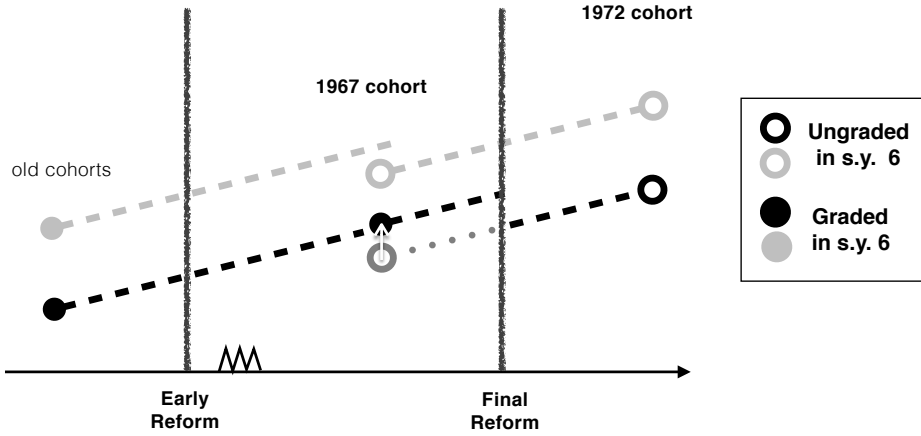


Figure 1.8: Difference in differences identification strategy

(*parallel trends assumption*), it is possible to isolate the effect of early grades. This situation is pictured in Figure 1.8: while the two sets of municipalities exhibit differences in outcomes unrelated to grade assignment, these differences are stable over time. Observed outcomes for the 1967 treated cohort can be compared to the counterfactual outcomes that would have been observed for the same set of municipalities absent the treatment (early grades). This counterfactual is given by the trend observed for the ungraded municipalities, assumed to be the same for treated municipalities. The effect of early grading is represented in the picture by the white arrow. The empirical specification that implements the difference in differences identification strategy is the following:

$$Y_{imc} = \alpha + \beta_{as} \text{Graded}_m \times 1967_c + 1967_c + \text{Munic}_m + \Delta X_{imc} + \epsilon_{imc} \quad (1.11)$$

$$a \in \{\text{Low ability, High ability}\}; s \in \{\text{Low SES, High SES}\},$$

where  $i$  indexes the individual,  $m$  the municipality, and  $c$  the cohort.  $\text{Munic}_m$  is a vector of fixed effects that captures persistent cross-sectional differences between municipalities.  $1967_c$  is a dummy that controls for the trend in outcomes. The variable  $\text{Graded}_m \times 1967_c$  picks up any differences in outcomes between grading and non-grading municipalities, that are not persistent, or the same, over time. Under the parallel trends assumption  $\beta_{as}$  represents the causal effect of early grading. Consistently with the model, the effect is allowed to differ by ability and SES, indexed respectively by  $a$  and  $s$  in equation 1.11. SES and ability are measured respectively using parental education and ability tests administered

in school year 6.<sup>24</sup> Notice that any determinant of the outcome that changes over time in a different way between the two sets of municipalities will also enter  $\beta_{as}$ , and thus bias the coefficient. Observable compositional change can be controlled for in the regression by adding  $X_{imc}$ , a vector of time varying pre-treatment controls. These covariates also increase precision of the estimates.

### 1.5.2 Inference

Sample size is large (around 18000 observation), but the treatment, grade assignment, varies at the municipal level. There are 29 municipalities in my sample, and half of them are treated before the reform. I conservatively cluster standard errors at the municipal level, rather than at the municipal-cohort level, which would result in twice as many clusters.<sup>25</sup> While the standard solution is to use cluster robust standard errors (Arellano, 1987; White, 1984), the number of clusters must be high for these standard errors to be unbiased. Cameron et al (2008) show that cluster-robust standard errors are downward biased in samples with few balanced (equally sized) clusters. They instead propose to use Cluster Bootstrap-t methods with null hypothesis imposed, and find that these methods yield the right p-values even with relatively few clusters (as few as 20). In a recent working paper MacKinnon & Webb (2014) confirm the good performance of the Cluster Bootstrap-t in the realistic case in which clusters are unbalanced. The Cluster Bootstrap-t is shown to perform well when treatment has enough variance.

My sample consists of 29 municipalities, both small and big. Treatment is given by the interaction between belonging to the cohort born 1967 and studying in an early grading municipality, which holds for about a quarter of the sample. There are thus enough clusters and treatment variation to believe that the Cluster Bootstrap-t should guarantee unbiased standard errors in my analysis. So in all my specifications I bootstrap standard errors using the method suggested by Cameron et al (2008). I also use sample weights to recover nationally representative estimates.

---

<sup>24</sup>Appendix 1.B.1 provides further details on ability and SES measures, and on the way I discretize them to match the model.

<sup>25</sup>This is suggested in Bertrand, Duflo, and Mullainathan (2004) for the case of panels. My final dataset is instead a cluster-panel, so there should be less correlation between clusters over time.

### 1.5.3 Testing for Identifying Assumptions

Difference in differences identifies the causal effect of assigning early grades under a specific set of assumptions. The most important one, as discussed before, is the parallel trends assumption: outcomes should trend similarly in both early grading and late grading municipalities. The assumption is more credible when the treated and untreated populations are not so different, especially in terms of “characteristics that are thought to be associated with the dynamics of the outcome variable” (Abadie 2005). This was shown to be the case above. In Appendix 1.C.1 I use administrative data from Statistics Sweden to test whether education and its determinants evolve in the same way in the two sets of municipalities: all tests pass. In particular trends in education for cohorts who went through compulsory school when all municipalities had abolished early grades (cohorts born 1969 onwards) appear to be parallel. The evidence thus supports the main assumption underlying the identification strategy.

A testable assumption of the identification strategy is that differences between treatment and control group in determinants of the outcome should be stable over time (e.g., there should be no compositional change). In the same way, response rates should be the same between treated and controls units over time (e.g., there should be no differential attrition).<sup>26</sup> In Appendix 1.C.2 I test for differential attrition and compositional change in the sample. First, there is no differential response to the student surveys and, importantly, I find no differential attrition in availability of SES and ability data. Second, it appears that the cross-sectional differences between grading non-grading municipalities are broadly stable over time. I find compositional change in specific parental occupations and education levels. Therefore in my final specification I also control for occupational dummies and parental education.<sup>27</sup>

A further assumption in the difference in differences setup is that the treated population should not change as a reaction to treatment assignment. In my setup this means that the students born 1967 should not enroll into different schools to get/avoid early grades. As catchment areas determined the compulsory school the student attended, parents had to relocate to a different municipality

---

<sup>26</sup>Both compositional change and differential attrition can lead to biased difference in differences coefficients (Blundell & Costa Dias, 2009).

<sup>27</sup>Results are robust to excluding income and parental controls. I include those variables to increase precision of the estimates.

if they desired a different grading policy for their children. Alternatively they could send their children to a private school. The first scenario seems highly unlikely, while private schools were not common in that period.

Finally it is important for identification that treatment and control group do not undergo different shocks over time. The presence of concurrent education reforms would be a problem in my setup if they were implemented at the municipal level. During the period I consider, schooling was quite centralized, with *national curricula* determining most of school policies. There is thus little scope for additional policies being differentially implemented in the two sets of municipalities. On top of that, the two cohorts I use in my analysis received their education in a relatively stable educational system: Sweden had already implemented the reforms of the 60s for the 9-year inclusive compulsory school, while the market-oriented school reforms of the 90s did not affect these cohorts.<sup>28</sup>

## 1.6 Empirical Results

The outcomes in the empirical analysis match those of the model. This allows me to understand whether empirical findings are consistent with students learning about their academic ability from grades. I thus investigate the effect of early grades on short-term effort choices, high school choices and attainment, and, finally, educational attainment and income. I also consider an alternative mechanism through which grades might affect education choices: grades might motivate/demotivate students, and thus affect their welfare.

I present difference in differences estimates from specification 1.11, which I re-parametrize to directly get coefficients for each ability – SES cell. In all specifications I control for ability (verbal and inductive ability normalized to the cohort-treatment level), basic demographics (gender, birth year, foreign status, special education), SES (income, parental occupation dummies, and education) and school-level variables (class size and teacher changes). For every outcome I report the point estimate, the p-value in parentheses, and, as a reference, the sample mean in brackets.<sup>29</sup>

---

<sup>28</sup>The reforms are described respectively by Meghir & Palme (2005) and Björklund et al (2005).

<sup>29</sup>The wild cluster bootstrap with null imposed does not yield standard errors.

There are two caveats when interpreting results. First, estimates are not very precise, so I can not detect very small effects. Second, I test many hypotheses, which in principle creates problems of false null rejection. Notice that the two problems go in opposite directions, and that the multiple hypothesis testing problem is less severe than it seems: most of the outcomes are strongly correlated, or can be considered different proxies for the same underlying variable (e.g., grades and course choices proxy for effort choice). Keeping this in mind, when I interpret results I focus on the overall picture rather than on single coefficients.

### 1.6.1 Effort in Compulsory School

In Tables 1.6 and 1.7 I investigate effects of early grades on school effort. The first Table reports effects on math and English course choices, which can be interpreted both as effort choices (academic courses are more challenging), and as early school choices reflecting future track selection (advanced courses are good preparation for academic high school). The second Table reports effects on grades in late compulsory school, which are straightforward measures of school effort.<sup>30</sup>

Low-ability students, especially those with low-SES, reacted to early grade assignment by switching to non-academic English, which can be interpreted as a reduction in effort (columns 2 and 3 in Table 1.6). The switches appear in grade 8, the first time in which the students could respond to grades released at the end of school year 6, and persist in school year 9.<sup>31</sup> Switches in course choice are observed for English, but not for math. One possible explanation is that parents and children already had feedback in math due to the correction of exercises. At this proficiency level parents could probably test children's math skills more than their English proficiency.<sup>32</sup> High-ability students did not revise course choices when graded early on.

Low-ability low-SES students exhibit worse math performance when graded early on (see column 2 and 3 of Table 1.7). High-ability high-SES students

---

<sup>30</sup>This is especially true of Swedish, a subject that does not involve any additional choice.

<sup>31</sup>The courses were chosen at the end of school year 6 for year 7, before final grades were released.

<sup>32</sup>The parents of the treated students were born in the 40s: at that time English proficiency was less widespread among parents than it is now the case in Sweden.

show instead higher English and Swedish grades when they receive the early grades. One can clearly see from the standardized Swedish test, which has more variation due to the different scale, that all low-ability students performed worse after being assigned early grades, while high-ability students performed better. Negative effects are stronger for low-SES students, positive effects are instead more pronounced for the high-SES students. In the aggregate no effect is found, as both positive and negative effects are summed up. This confirms the importance of looking at heterogeneous effects. The pattern found in the model is thus reproduced by the data: low (high) ability students are putting less (more) effort, and effects are stronger for low (high) SES students. However, high-ability high-SES students are putting more effort, rather than reducing it, as in the model. This implies that “*substitution*”, rather than “*income effects*”, are prevailing. This can be easily rationalized within the model, assuming that different majors require different ability levels. Then it is easy to see that these students would react to high grades by further increasing effort.

Table 1.6: Effects on course choices (school years 7-9):  
Summary of difference in differences estimates

Outcome:	Low-ability			High-ability	
	All Sample	Low-SES	High-SES	Low-SES	High-SES
<i>Advanced Math</i> (school year 7)	0.00 (0.94) [0.73]	-0.03 (0.70) [0.54]	0.01 (0.84) [0.72]	0.04 (0.50) [0.90]	0.04 (0.47) [0.95]
<i>Advanced Math</i> (school year 8)	-0.01 (0.80) [0.66]	-0.00 (1.00) [0.43]	-0.04 (0.26) [0.64]	0.02 (0.40) [0.87]	-0.00 (0.94) [0.96]
<i>Advanced Math</i> (school year 9)	0.02 (0.64) [0.57]	0.01 (0.74) [0.32]	-0.00 (0.97) [0.53]	0.03 (0.58) [0.76]	0.05 (0.16) [0.90]
<i>Advanced English</i> (school year 7)	0.00 (0.90) [0.75]	-0.03 (0.60) [0.57]	-0.01 (0.88) [0.76]	0.05 (0.15) [0.91]	0.05 (0.21) [0.97]
<i>Advanced English</i> (school year 8)	-0.05** (0.02) [0.73]	-0.06*** (0.01) [0.53]	-0.07 (0.19) [0.73]	-0.01 (0.48) [0.91]	-0.01 (0.63) [0.97]
<i>Advanced English</i> (school year 9)	-0.06* (0.06) [0.68]	-0.07** (0.02) [0.46]	-0.08 (0.12) [0.65]	-0.05 (0.14) [0.87]	-0.01 (0.58) [0.95]

\*  $p < 0.10$ , \*\*  $p < 0.05$ , \*\*\*  $p < 0.01$

Wild Cluster Bootstrap p-values in parentheses; sample averages in brackets. All specifications control for basic demographics, relative ability measures (standardized at the treatment-cohort level) and parental background.



Table 1.7: Effects on grades (school years 8 and 9):  
Summary of difference in differences estimates

Outcome:	Low-ability			High-ability	
	All Sample	Low-SES	High-SES	Low-SES	High-SES
<i>Math Grade</i> (school year 8)	-0.02 (0.75) [3.04]	-0.08* (0.07) [2.70]	-0.04 (0.64) [2.87]	0.05 (0.14) [3.35]	0.07 (0.32) [3.59]
<i>Math Grade</i> (school year 9)	-0.11 (0.12) [3.20]	-0.15** (0.02) [2.86]	-0.13 (0.20) [3.03]	-0.04 (0.68) [3.54]	-0.06 (0.42) [3.73]
<i>English Grade</i> (school year 8)	0.06 (0.25) [3.05]	0.00 (0.98) [2.70]	0.10 (0.11) [2.86]	0.02 (0.79) [3.37]	0.17*** (0.00) [3.61]
<i>English Grade</i> (school year 9)	0.05 (0.21) [3.18]	-0.03 (0.57) [2.82]	0.12 (0.11) [3.05]	0.05 (0.48) [3.46]	0.16*** (0.00) [3.73]
<i>Swedish Grade</i> (school year 8)	0.03 (0.47) [3.06]	-0.04 (0.38) [2.64]	-0.03 (0.58) [2.91]	0.12 (0.13) [3.43]	0.17*** (0.00) [3.68]
<i>Swedish Grade</i> (school year 9)	0.06 (0.17) [3.17]	-0.02 (0.76) [2.69]	0.03 (0.64) [3.03]	0.14** (0.01) [3.54]	0.18*** (0.01) [3.86]
<i>Swedish Test</i> (school year 9)	0.18 (0.84) [15.84]	-3.64*** (0.00) [13.43]	-1.54** (0.01) [13.25]	4.49*** (0.00) [19.96]	6.14*** (0.00) [18.51]

\*  $p < 0.10$ , \*\*  $p < 0.05$ , \*\*\*  $p < 0.01$

Wild Cluster Bootstrap p-values in parentheses; sample averages in brackets. Math and English pool together grades for advanced and general courses. All specifications control for basic demographics, relative ability measures (standardized at the treatment-cohort level) and parental background.

## 1.6.2 Education and Income

In Table 1.8 I report effects of early grades on high school choices, educational attainment, and income. Contrary to what the model predicts, early grades do not lead to different high school track choices. I observe instead an increase in enrollment for all students. While this can be surprising (on average low-ability students reduced effort in compulsory school), it is possible that lowest ability students increased effort early on, and thus decided to enroll into high school. This “*income effect*” was discussed in the model in Section 1.4.

When looking at educational attainment, I find an increase in high school attainment at age 17-20 for high-ability low-SES students, mostly explained by a reduction in high school dropout. In the long-run this effect becomes smaller and close to insignificant. In Sweden adult education programs (*Komvux*) allow people to complete further education: in the counterfactual scenario of late grading students might still have been able to finish their high school education. Moreover, I find that low-ability low-SES students are less likely to attain college. These effects are qualitatively consistent with model’s predictions: a reduction in dropout due to higher effort in compulsory school, and less low-ability students ending up with an academic education.

Why do the short-run effects of early grades do not pass on to high school track choice, and why is educational attainment not affected for high SES-students? I propose as an explanation that preferences for education might attenuate the effects of early grades. In Appendix 1.B.2 I show that, controlling for ability, academic high school enrollment rates of high-SES students are 20 percentage points higher than those of low-SES students. At the same time grade differences in late compulsory school between high- and low-SES students are at most  $\frac{1}{4}$ th of a grade. SES appears thus to strongly influence high school choices in Sweden, independently of ability.

While it is important to assess how early grades affect education outcomes to understand mechanisms, a full evaluation of the policy requires looking at long-run outcomes. Early grade assignment does not significantly affect income at ages 33-40, a good proxy of lifetime income in the Sweden labor market (Börklund, 1993). This is consistent with the theoretical model, which also generated very small effects on lifetime income. Early grading leads instead to an increase in upward income mobility among low-ability low-SES students,

who displayed the strongest downward revisions in education choices.<sup>33</sup> I conclude that, from the perspective of the labor market, early grades simply allowed students to better sort by ability into education. For low-ability low-SES students this implies a reduction of over-investment in education, and potentially an increase in total earnings.

---

<sup>33</sup>I consider upward mobile a student if she is 15 percentile ranks above the parents' income percentile rank.

Table 1.8:  
Effects on high school choices, educational attainment and income:  
Summary of difference in differences estimates

Outcome:	Low-ability			High-ability	
	All Sample	Low-SES	High-SES	Low-SES	High-SES
<i>HS Enrollment</i> (age 15-18)	0.04** (0.02) [0.89]	0.03* (0.08) [0.85]	0.06** (0.03) [0.92]	0.03* (0.06) [0.93]	0.03** (0.03) [0.97]
<i>Academic HS Track</i> (age 15-18)	0.02 (0.55) [0.47]	0.01 (0.82) [0.20]	0.01 (0.85) [0.44]	0.02 (0.68) [0.59]	0.04 (0.13) [0.81]
<i>HS Dropout</i> (age 17-20)	0.00 (0.98) [0.13]	0.02 (0.46) [0.18]	-0.01 (0.53) [0.12]	-0.05** (0.02) [0.10]	0.02 (0.34) [0.08]
<i>Attains HS</i> (age 17-20)	0.00 (0.79) [0.79]	-0.02 (0.49) [0.72]	0.02 (0.35) [0.83]	0.06*** (0.01) [0.86]	-0.01 (0.82) [0.91]
<i>Attains HS</i> (age 33-40)	0.00 (0.94) [0.92]	-0.02 (0.19) [0.88]	0.03 (0.10) [0.95]	0.02 (0.14) [0.96]	-0.00 (0.87) [0.98]
<i>College or more</i> (age 33-40)	-0.02 (0.28) [0.43]	-0.03* (0.06) [0.22]	0.02 (0.59) [0.42]	-0.04 (0.23) [0.52]	0.00 (0.94) [0.75]
<i>Gross income</i> (age 33-40)	3.28 (0.61) [259.11]	11.49 (0.21) [223.88]	-3.64 (0.75) [256.31]	-6.76 (0.62) [269.89]	0.78 (0.95) [330.13]
↑ <i>Income mobility</i> (age 33-40)	0.04** (0.02) [0.34]	0.08*** (0.00) [0.38]	0.02 (0.54) [0.27]	0.01 (0.60) [0.45]	0.02 (0.60) [0.28]

\*  $p < 0.10$ , \*\*  $p < 0.05$ , \*\*\*  $p < 0.01$

Wild Cluster Bootstrap p-values in parentheses; sample averages in brackets. HS Enrollment is measured at ages 16-18, HS attainment at age 40. Income is measured at ages 33-40. ↑ Income mobility is 1 when student income is 15 ranks above family income rank. All specifications control for basic demographics, relative ability measures (standardized at the treatment-cohort level) and parental background.

### 1.6.3 Student Welfare

Part of the policy debate in Sweden, and in other countries that considered early grades abolition, revolved around the concern that grades might demotivate (motivate) students who put low (high) effort independently of their ability, and create a competitive environment where weak students fare worse. If this was the case students would derive disutility (utility) from low (high) grades, and their preferences for education would be affected by their performance.

To validate this alternative mechanism I investigate the effects of early grading on self-reported child welfare. Outcomes are taken from the student surveys. The first survey was administered in school year 6, before final grades were assigned. It should pick up potential effects due to the more competitive/challenging environment. The second survey was assigned in school year 10, and asked many retrospective questions about how children were feeling in late compulsory school, when I observe most of the effects of early grades. Tables 1.9 and 1.10 show that, all in all, early grades did not significantly affect student welfare.<sup>34</sup> The only statistically significant effects are found for low-ability low-SES students, who are less likely to report that they do well in school before getting the grades, and also are less likely to report that they enjoyed late compulsory school (school years 7-9). While the first finding is not negative per se, since it shows that these students were more conscious of their school performance, the second one might be more concerning for policy-makers. However similar outcomes pertaining to school welfare show a 0 effect also for these students, so I am more inclined to consider the finding a spurious effect.

---

<sup>34</sup>As explained before I cannot detect small effects, but I can state that there appears to be no major effect.

Table 1.9: Effects on behavior in school year 6:  
Summary of difference in differences estimates

Outcome:	Low-ability			High-ability	
	All Sample	Low-SES	High-SES	Low-SES	High-SES
<i>I do well in school</i>	-0.04 (0.15) [0.73]	-0.07** (0.02) [0.61]	-0.08 (0.13) [0.69]	-0.01 (0.71) [0.84]	0.03 (0.23) [0.89]
<i>Parents think I do well in school</i>	-0.02 (0.27) [0.89]	-0.02 (0.17) [0.84]	-0.06 (0.11) [0.86]	-0.03 (0.19) [0.94]	0.03 (0.11) [0.96]
<i>I do my best, even if boring</i>	-0.02 (0.30) [0.71]	-0.03 (0.29) [0.73]	-0.03 (0.16) [0.71]	-0.01 (0.77) [0.71]	-0.01 (0.88) [0.67]
<i>I want to improve in school</i>	-0.00 (0.91) [0.59]	-0.01 (0.77) [0.71]	-0.02 (0.53) [0.66]	0.04 (0.36) [0.47]	-0.03 (0.51) [0.44]
<i>I dislike answering questions</i>	-0.02 (0.34) [0.16]	-0.03 (0.21) [0.21]	-0.01 (0.79) [0.18]	0.00 (0.87) [0.14]	-0.04* (0.07) [0.10]
<i>I learn useless stuff at school</i>	0.01 (0.69) [0.38]	0.03 (0.33) [0.39]	0.03 (0.37) [0.38]	-0.00 (0.97) [0.40]	-0.01 (0.80) [0.34]
<i>I get disappointed if I get bad scores</i>	0.01 (0.65) [0.68]	-0.02 (0.58) [0.63]	0.02 (0.33) [0.70]	0.04 (0.34) [0.69]	0.03 (0.26) [0.75]

\*  $p < 0.10$ , \*\*  $p < 0.05$ , \*\*\*  $p < 0.01$

Wild Cluster Bootstrap p-values in parentheses; sample averages in brackets. All specifications control for basic demographics, relative ability measures (standardized at the treatment-cohort level) and parental background.

Table 1.10: Effects on behavior in late compulsory school:  
Summary of difference in differences estimates

Outcome:	Low-ability			High-ability	
	All Sample	Low-SES	High-SES	Low-SES	High-SES
<i>I enjoyed grades 7-9</i>	-0.01 (0.41) [0.72]	-0.03* (0.07) [0.71]	-0.01 (0.65) [0.71]	-0.01 (0.69) [0.75]	0.01 (0.67) [0.74]
<i>I was worried in grades 7-9</i>	0.01 (0.79) [0.12]	-0.01 (0.72) [0.13]	0.05 (0.65) [0.10]	0.01 (0.75) [0.11]	-0.00 (0.98) [0.12]
<i>I am happy with grades 7-9</i>	0.00 (0.81) [0.75]	-0.01 (0.89) [0.68]	0.00 (0.86) [0.71]	0.02 (0.65) [0.82]	0.00 (0.86) [0.85]
<i>I got help at home in grades 7-9</i>	-0.01 (0.66) [0.71]	-0.01 (0.52) [0.67]	-0.02 (0.30) [0.75]	-0.01 (0.62) [0.68]	0.02 (0.33) [0.76]
<i>I did my best even if boring</i>	0.01 (0.62) [0.47]	-0.00 (0.93) [0.50]	0.05 (0.14) [0.47]	0.02 (0.28) [0.47]	-0.00 (1.00) [0.45]
<i>I did my best even if hard</i>	-0.00 (1.00) [0.71]	-0.01 (0.44) [0.71]	0.04 (0.17) [0.70]	-0.02 (0.26) [0.71]	0.02 (0.55) [0.74]
<i>I learned useless stuff at school</i>	0.03 (0.19) [0.53]	0.04 (0.27) [0.57]	0.03 (0.19) [0.52]	0.05 (0.17) [0.54]	0.01 (0.84) [0.47]
<i>I was stressed at school</i>	0.01 (0.65) [0.20]	0.01 (0.76) [0.20]	0.00 (0.86) [0.22]	0.00 (0.93) [0.19]	0.02 (0.39) [0.19]

\*  $p < 0.10$ , \*\*  $p < 0.05$ , \*\*\*  $p < 0.01$ 

Wild Cluster Bootstrap p-values in parentheses; sample averages in brackets. All specifications control for basic demographics, relative ability measures (standardized at the treatment-cohort level) and parental background.

## 1.7 Discussion

In the following I assess how my results compare to previous literature on the effects of grading information. Stinebrickner & Stinebrickner (2012) study dropout behavior in Berea college, an institution with free tuition and subsidized boarding catering to disadvantaged students. They find that dropout is strongly explained by students revising downward their priors on academic performance. Similarly Zafar (2011) finds that Northwestern undergraduates revise downward their beliefs, and switch to easier majors, when they observe grades lower than predicted. In Zafar's paper the deviation between expected and realized academic performance is taken as an "information metric" that identifies new information about students' "own unobserved academic ability." In fact this information might reflect, as explicitly recognized in Stinebrickner's paper, college preparation rather than academic ability. In both studies it is not possible to determine whether the updates are on academic ability or the stock of knowledge accumulated. While both problems signal the need for better selection into college, from the policy perspective they have quite different implications.<sup>35</sup>

The results of my paper are in line with the learning mechanism outlined by Stinebrickner & Stinebrickner (2012) and Zafar (2011) at the college level. In particular, the responses I find along the ability distribution are consistent with students revising their priors about ability. In my setup grades were assigned when children were 13 years old, so there is less concern that the update is on previous preparation, rather than ability. On top of that I show that students with the same SES (which could proxy for early effort), but different ability levels, react differently to grade assignment. This is consistent with students learning about ability rather than previous preparation. My paper shows both theoretically and empirically that the reaction to grades differs by SES. In the above-mentioned two papers there is no variation in SES, as sampled students are either low-SES (Berea college) or high-SES (Northwestern undergraduate students).

My paper is also related to the grading standards literature, which stresses the role of ability in students' responses to grades. Becker & Rosen (1992) and

---

<sup>35</sup>Failure in evaluating own ability calls for a revision of grading information. Failure in college preparation requires to revise curricula in earlier education tiers.



Betts (1998) show theoretically that higher grading standards encourage high ability students to put more effort, while students below standard might be discouraged. Betts & Grogger (2003) empirically confirm the heterogeneous effects of increasing grading standard at the high school level, while Figlio & Lucas (2004) find that higher standards lead to positive results on test scores, with effects that depend on the ability of the student relative to the class. In my setup untreated students do not observe grades, but only test scores. Absent grades, low-SES students are likely to have lower grading standards than high-SES students (for instance because the difficulty of the tests follow class ability). Introducing grades should thus lead to positive effects for high-SES students and negative effects for low-SES students. My results do not confirm this, and are rather consistent with students learning about their ability from grades.

The grading reform I analyze has been previously studied by Sjögren (2010), who uses administrative data to study long-run effects (final education and income) of the overall reform using difference in differences. She finds evidence of a positive effect of early grading on educational attainment for girls, and a negative effect for high-SES students. Differences in educational attainment are found also before and after the reform took place, which casts some doubts on the robustness of the results. My paper focuses on the mechanisms through which grades affect education choice: the theoretical model I develop shows that average treatment effects could mask substantial heterogeneity in the response to grades, which suggests conditioning the analysis at least by ability. I do not find any negative effect of early grades for high-SES students, while I do not study effects by gender, as this would exacerbate the multiple hypothesis testing problem. Results appear to be more robust, as trends in educational attainment for treated and control municipalities appear to be parallel in the refutability tests. This is likely due to the different cohorts used: Sjögren's sample comprises twenty cohorts, so she needs to assume parallel trends over two decades. My sample uses cohorts who studied just before and after the final reform. I only need to assume that trends between the two sets of municipalities are the same within a 5-year window, which I show to be the case.

## 1.8 Conclusion

In this paper I investigate the effect of early grades on students' education choices and attainment. I exploit the staggered implementation of a curriculum reform,

which postponed grade assignment in Swedish compulsory schools, to estimate both short- and long-run effects of early grading. To investigate mechanisms I compare empirical results to the predictions of a sequential choice learning model based on the setup.

In the model children are uncertain about academic ability, and their priors differ by socioeconomic status (SES). Grades are ability signals that allow children to re-optimize educational choices. The calibrated model shows that early grading results into choices closer to first best for all students: low-ability students reduce effort in compulsory school and are more likely to choose vocational high school. High-ability students increase effort in school and are more likely to choose academic education paths. Stronger responses are found for students who observe information consistent with their priors, so that effects differ by SES.

The empirical results of my analysis are in line with the theoretical predictions for effort choices in late compulsory school. When graded early on, low-ability low-SES students are more likely to get lower grades and switch to easier courses in compulsory school than high-SES students with similar levels of ability. High-ability students, especially if high-SES, are more likely to get higher grades in late compulsory school when graded early on. Contrary to what the model predicts, early grades do not affect high school track choices and educational attainment for high-SES students. I find that high school attainment increases by 6 pp for high-ability low-SES students, while college attainment decreases by 3 pp for low-ability low-SES students. What explains the differences between model predictions and empirical findings at the high school level? The data suggests that SES strongly influences high school choices in Sweden, independently of ability. This might attenuate the effects of early grades. None of the effects found on education carry over to the labor market. In particular I find no effects on lifetime income, measured at ages 33-40. This suggests that early grading information simply improved the match between early education choices and ability, and reduced over-investment in education. Finally I find no evidence of demotivating effects for low-SES students, one of the main concerns that motivated the grading reforms.

The key economic implication of my results is that students are uncertain about their ability in early stages of education, when I show that grades affect their choices. This contrasts with the workhorse models of education choice (Becker, 1994; Ben-Porath, 1967), that assume complete information and thus

no ex-ante uncertainty in the returns to non-compulsory education. From the policy point of view, I establish that early grading leads to a better match between education and ability, but increases inequality in educational attainment and reduces effort in compulsory school for low-ability students. Whether early grading is a desirable policy depends thus on the objective function of the policy-maker.

With regard to future research directions, it is possible to expand the scope of the analysis by looking at further sources of heterogeneity, which relate to different mechanisms. First, as whole classes are sampled in my data, I can look at differential effects of grading depending on relative ability. If students judge ability against their immediate peers, average ability students in high-ability (low-ability) classes might react more positively (negatively) to early grades. Second, average ability students might get more information out of grades, as it is less likely that they get top or bottom scores in homework and tests.<sup>36</sup> Third, there could be different responses to grading information along the gender dimension, as boys are found to generally be more overconfident than girls in ability (Bertrand, 2011). Before exploring these additional sources of heterogeneity it is however important to correct standard errors for multiple hypothesis testing. This requires some additional work, as standard errors are already bootstrapped to deal with the small number of clusters.

Lastly, it is worthy to investigate theoretically alternative mechanisms through which grades affect education choices, including the ones outlined above. For instance what happens when ability signals reflect knowledge rather than ability? This can be the case if parents and students are not able to distinguish the ability component of the grade from previous effort choices. I argue this is a plausible mechanism when parents do not observe children's effort, but do observe the final grades.

---

<sup>36</sup>Relatedly, Stange (2012) reports that the students for whom college grades, and the option to dropout in college, have the highest value, are moderate ability students, who have the strongest uncertainty about finishing college.

## 1.A Numerical Model

### 1.A.1 Evidence on assumptions and calibration

The model makes precise assumptions about choice protocols, the distribution of ability in the population, school selection, and payoffs to education. In this section I provide evidence supporting model assumptions, and discuss calibration.

The basic assumption underlying the model is that students are forward looking in the education choices. Table 1.A.1 reports summary statistics on the items that surveyed students considered important when choosing high school. Apart from preferences for the chosen program, the items that rank highest are study plans, ability and grades. This shows that students were forward-looking in their choices, and considered feasibility of the chosen track important in their choices.

Table 1.A.1: Survey evidence on HS choice, 1967 cohort

	Mean	Obs
Chose HS after interest	0.80	6195
Chose HS after study plans	0.64	6099
Chose HS after ability	0.60	6093
Chose HS after grades	0.49	6117
Chose HS after parents	0.23	6099
Chose HS after peers	0.07	6098

Data from grade 10 survey. All variables represent agreement with the statement and are coded from 0 to 1 (1 represents full agreement).

In the model high SES students are assumed to have higher levels of ability than low SES students. Figure 1.A.1 confirms this empirically.<sup>37</sup> While low-SES students have normal ability distributions, high-SES students display left-skewed distributions. In Table 1.A.2 I calibrate the data to the discrete distribution in column 1. The resulting distributions by SES are then used to simulate ability distributions in the model.

<sup>37</sup>This is consistent with the evidence on early differences in ability through the socioeconomic gradient shown by Cunha & Heckman (2009).

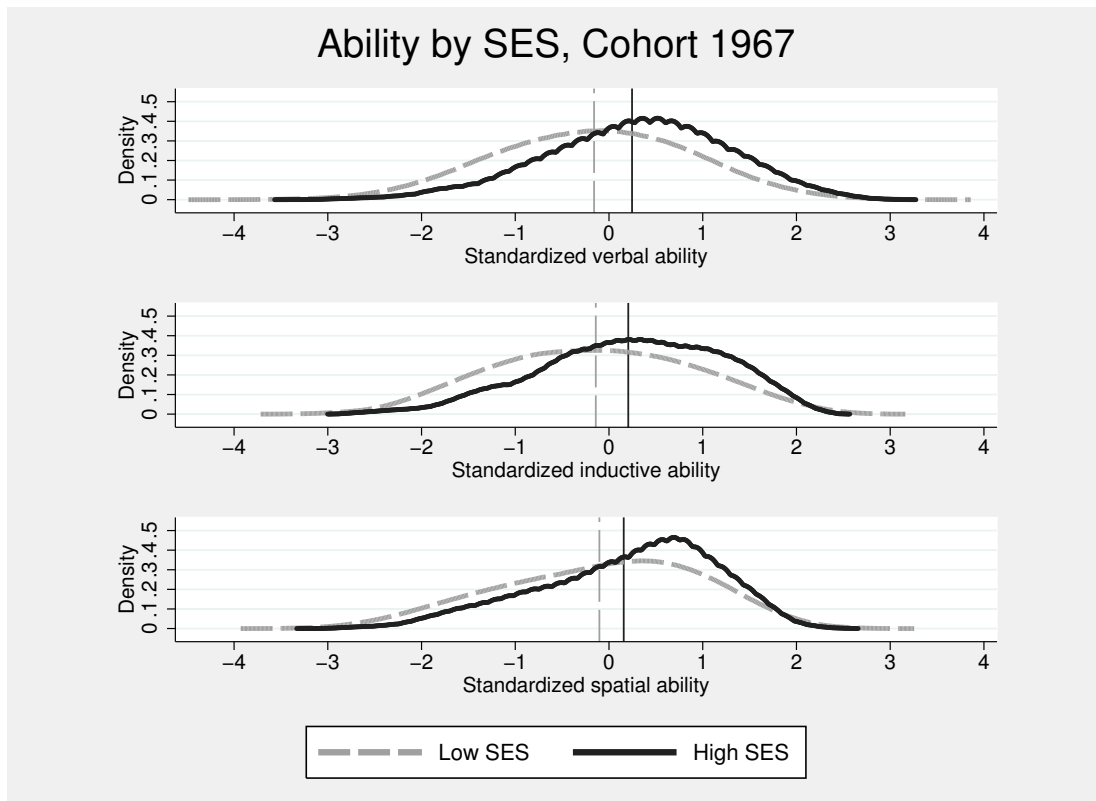


Figure 1.A.1: Differences in standardized ability by SES

Note: The SES division is based on parental education. Ability measures are taken from tests administered in school year 6, and are standardized at the treatment-cohort level.

Table 1.A.2: Distribution of discretized ability by SES, 1967 cohort

	All	Low SES	High SES
Lowest ability	0.10	0.13	0.06
Low ability	0.20	0.24	0.14
Medium ability	0.30	0.31	0.28
High ability	0.25	0.22	0.29
Highest ability	0.15	0.10	0.23

The population ability distribution is constrained to the bins in column 1. The distributions by SES are generated using the same cut-points.

Figure 1.A.2 shows ability levels by completed education. Students who attained high school have higher ability than those who dropped out of high school, or never enrolled. Students who have a college education have much higher ability levels, which is consistent with the assumptions I make in the model, and is not surprising given that college is highly selective.

Table 1.A.3: Income by Final Education,  
1967 Cohort

Completed education	Gross Income	Premium
Compulsory school	184.40	0.00
Vocational HS	221.21	0.20
Academic HS	226.22	0.23
College	290.67	0.58

Before-tax income measured at ages 33-40, in thousands kronor

Table 1.A.3 summarizes income premia for each education choice. While these are not causal estimates, they might be representative of the information that young students use when assessing their education goals. High school graduates exhibit higher incomes than students with compulsory education. As assumed in the model, the income of students with academic high school are not substantially different from those of students with vocational high school. The wages of college graduates<sup>38</sup> are instead quite higher.

In Figure 1.A.3 I plot the wages of the students in the sample by discretized ability. There is little variation in wages by ability for students with compulsory school or high school. However there seems to be complementarity between income and ability for college graduates. In the model I thus allow the wage premium for college to depend on ability, and use the estimates as payoffs.

<sup>38</sup>Here including also 2-year short college.

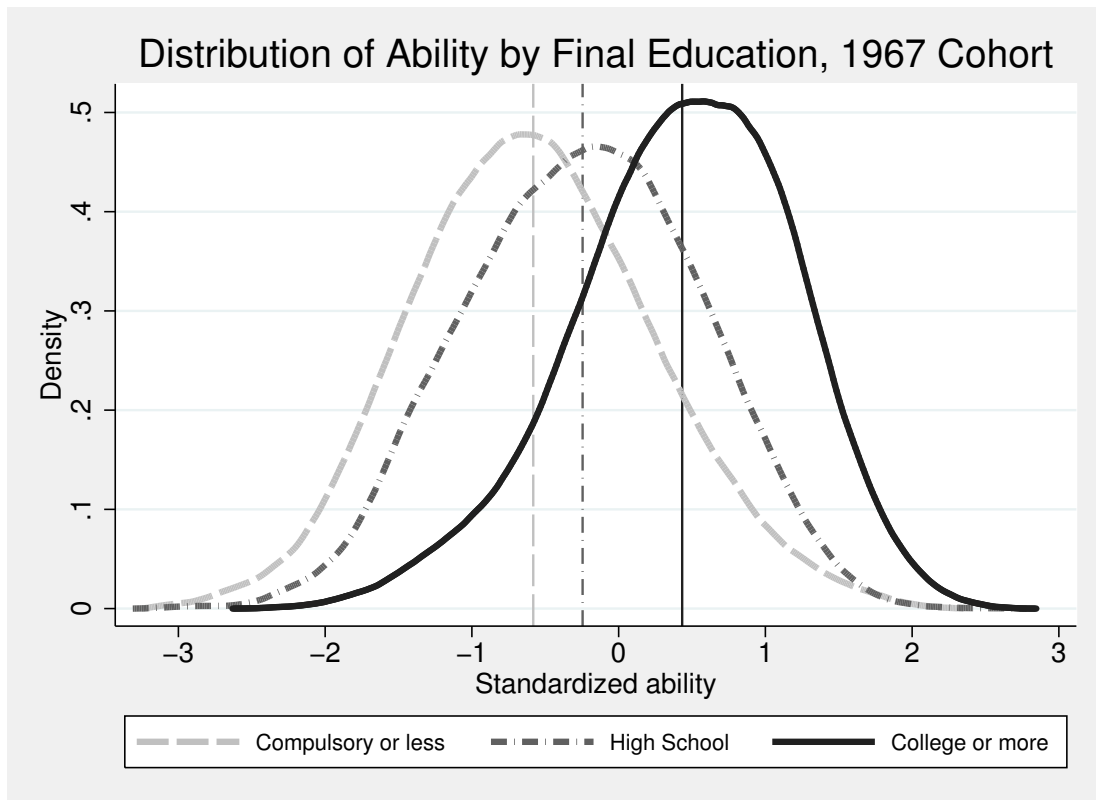


Figure 1.A.2: Standardized ability by final attained education.

Note: Ability measures are taken from tests administered in school year 6, and are standardized at the treatment-cohort level.

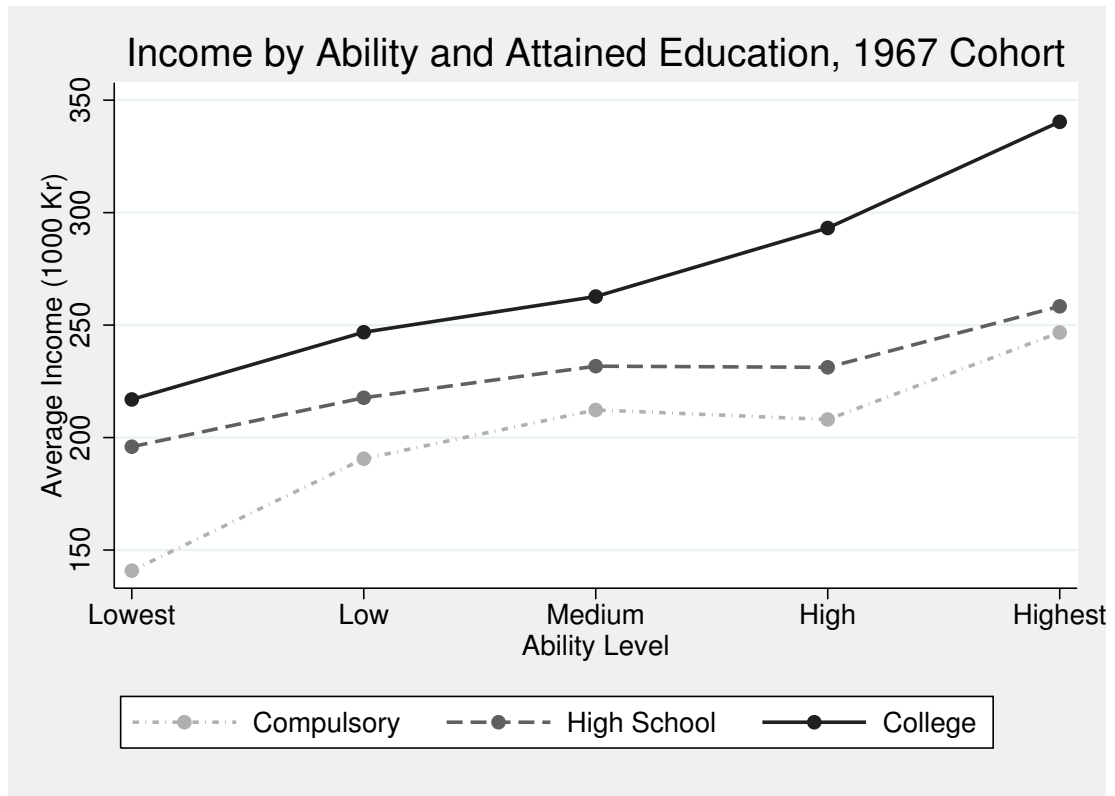


Figure 1.A.3: Average gross income by ability and attained education, 1967 cohort

Note: Before-tax income measured at ages 33-40, in thousands kronor. Ability is discretized to a 1-5 scale, as in the model.



I calibrate the knowledge production function,  $k_{it} = \omega_t(\alpha a_i + \beta e_{it}) + \delta k_{it-1}$ , using the following parameters:

Table 1.A.4: Production Function Parameters	
weights	coefficients
$\omega_1 = 4/9$	$\alpha = 1$
$\omega_2 = 1/9$	$\beta = 2.5$
$\omega_3 = 5/9$	$\gamma = 1.1$

Table 1.A.5 reports minimum effort and ability levels required to access and attain each education level. The knowledge thresholds are found substituting the values for each education level into equation 1.12:

$$k_3 = \alpha a \times (\omega_1 \gamma^2 + \omega_2 \gamma + \omega_3) + \beta(\omega_1 \gamma^2 \times e_1 + \omega_2 \gamma \times e_2 + \omega_3 \times e_3). \quad (1.12)$$

The value of vocational school,  $V_{i,t}^{E_2}$ , was shown in Section 1.3.1 to be:

$$V_{i,\tau}^{E_2} = \sum_{t=\tau}^3 -\omega_\tau \times \underline{\gamma}_E (e_{it,\tau}^{E_2\star})^{\bar{\gamma}_E} + P(\tilde{k}_{i3,\tau} \geq \bar{k}^{E_2}) \times U((L-2) \times w_2) + P(\tilde{k}_{i3,\tau} < \bar{k}^{E_2}) U((L-1) \times w_1). \quad (1.13)$$

The parameters I use for effort disutility are:  $\underline{\gamma}_E = 6$  and  $\bar{\gamma}^E = 1.6$ . The parameters for income utility,  $U(I) = \underline{\gamma}_I (I_i^{E_2\star})^{\bar{\gamma}_I}$ , are:  $\underline{\gamma}_I = 2.8$  and  $\bar{\gamma}^I = 0.9$ . Effort costs are thus convex, and income utility is concave.

Table 1.A.5: Minimum ability and effort for Educational Attainment (Knowledge Thresholds)

	$\bar{k}^{E_2}$	$\underline{k}^{E_3}$	$\bar{k}^{E_3}$	$\bar{k}^{E_4}$
$a$	2	3	3	4
$e_1$	Medium	Medium	Medium	Medium
$e_2$	Medium	Medium	Medium	Medium
$e_3$	Low	Low	Medium	High

In the model high-ability students are more likely to observe higher grades than low-ability students do when graded early on. I confirm this in Figure 1.A.4, where I plot grades at the end of school year 6 (the treatment) for treated students born 1967. The vertical black line represents the average grade for each SES: it could be considered the prior grade the student is expected to get, before information about ability is revealed. For low-SES students the average grade is closer to the mean for low-ability students. The opposite is true for high-SES students. This reflects the different composition in ability within SES.

Finally In the model I assume that grades are unbiased. In Table 1.A.6 I try to assess this empirically. While there is a strong relationship between the standardized test and the final grade (the coefficient is close to 1), it appears that SES has an independent positive effect on final grades, controlling for ability. This could be due to a positive bias towards wealthier students, but could also be related to the fact that high-SES students put more effort into schooling. Given that final grades corrected for discrepancies between yearly and test performance, it is still possible that they are unbiased. Notice that the magnitude of this higher bound effect is actually small: one child over ten/twenty gets a higher grade if categorized as high-SES with respect to a low-SES child.

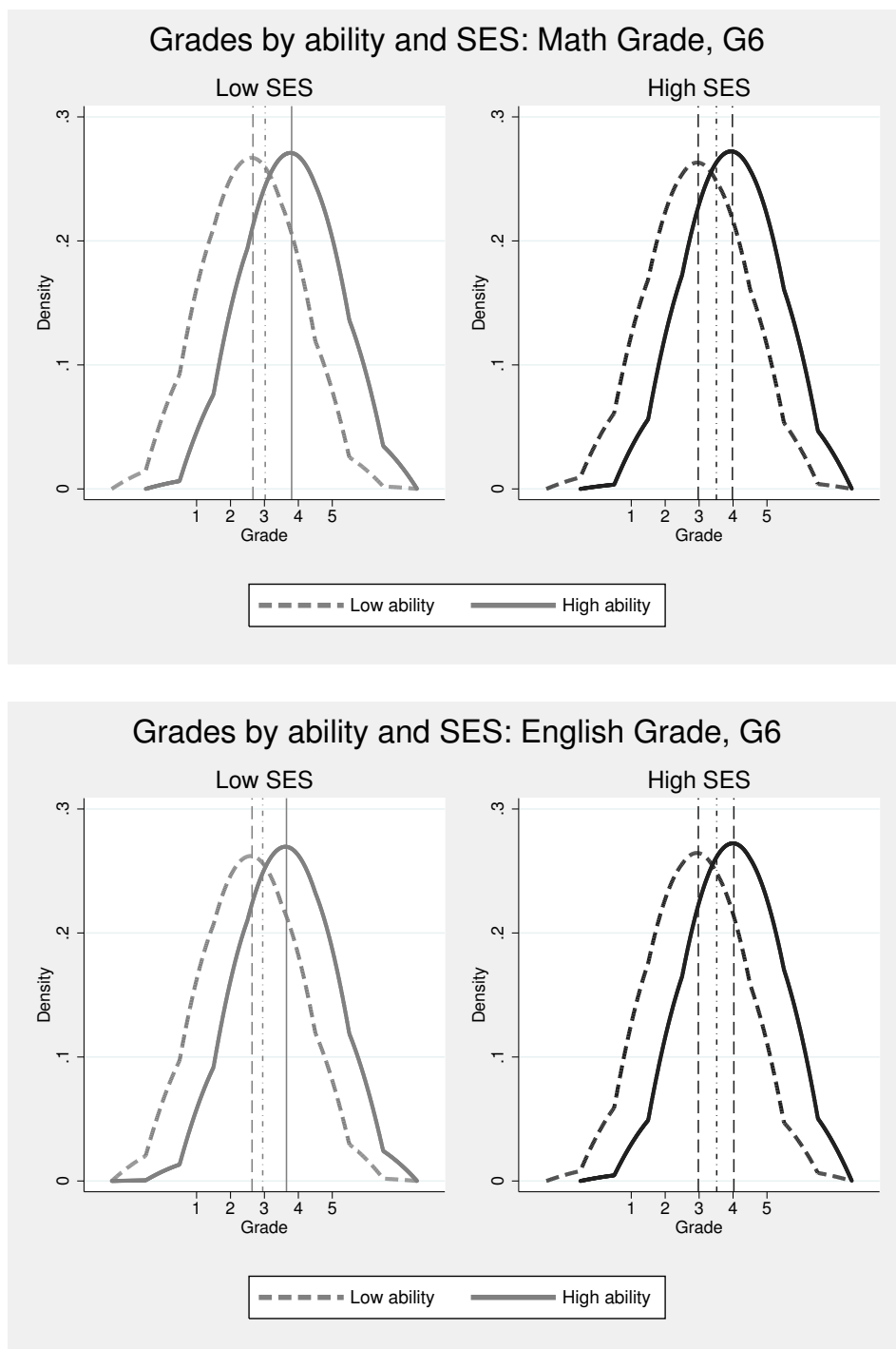


Figure 1.A.4: Grades by Ability and SES in school year 6

Note: Distributions over-smoothed for illustrational clarity. Dash-dot vertical lines represent averages for each SES cell. The other vertical lines represent averages for each ability–SES cell.

Table 1.A.6: Testing for biases in final grade assignment

	Adv Math grade (year 9)	Adv English grade (year 8)
<b>Regressor of interest:</b>		
High-SES	0.09*** (0.02)	0.04*** (0.01)
<b>Controls:</b>		
Normalized test score	0.69*** (0.01)	0.63*** (0.01)
Normalized ability	0.06*** (0.02)	0.11*** (0.01)
1967 cohort	-0.03 (0.02)	0.01 (0.02)
$R^2$	0.60	0.60
Observations	6535	8867

*Note:* \*  $p < 0.10$ , \*\*  $p < 0.05$ , \*\*\*  $p < 0.01$

Normalized test score and Ability refer to math tests in column 1, and English tests in column 2. The SES division is based on parental education. Standard errors are clustered at the class level.

## 1.A.2 Solution Method

**Simulation of ability signals**

I extract true ability and unbiased ability signals (See Figure 1.A.5) from a multivariate normal with covariance matrix:

$$\begin{bmatrix} 1 & 0 & 0 \\ 0 & \sigma_1^2(\epsilon) & 0 \\ 0 & 0 & \sigma_2^2(\epsilon) \end{bmatrix}$$

I discretize the normal draws using the SES-specific distributions shown in Table 1.A.2. I assume the following:

- The grade signals are a sum of true ability and noise:
  - $g_{i2} = a_i + \epsilon_2$  with  $\epsilon_2 \sim N(0, \sigma_2^2(\epsilon))$
  - $g_{i3} = a_i + \epsilon_3$  with  $\epsilon_3 \sim N(0, \sigma_3^2(\epsilon))$
  - $cov(a_i, \epsilon_2) = cov(a_i, \epsilon_3) = cov(\epsilon_2, \epsilon_3) = 0$
- Late grades are more precise than grades assigned in school year 6:
  - $corr(a_i, g_{i2}) = 0.7$
  - $corr(a_i, g_{i3}) = 0.8$

I need to find  $\sigma_1^2(\epsilon)$  and  $\sigma_2^2(\epsilon)$  such that  $k_2 = corr(a_i, g_{i2}) = 0.7$  and  $k_3 = corr(a_i, g_{i3}) = 0.8$ :  $k_t = corr(a_i, g_{it}) = corr(a_i, a_i + \epsilon_t) = \frac{1 + 0}{\sigma(a_i) + \sigma(a_i + \epsilon_t)} = \frac{1}{\sigma(a_i) \times \sigma(a_i + \epsilon_t)} = \frac{1}{\sigma(a_i + \epsilon_j)}$  and  $\sigma^2(a_i + \epsilon_t) = 1 + \sigma_t^2(\epsilon)$ . So  $k_t = \frac{1}{\sqrt{1 + \sigma_t^2(\epsilon)}}$ , thus  $\sigma_t^2(\epsilon) = \frac{1}{k_t^2} - 1$ . Because  $\sigma^2(a_i + \epsilon_t) = \frac{1}{k_t^2}$  it follows that  $corr(g_{i2}, g_{i3}) = \frac{cov(a_i + \epsilon_2, a_i + \epsilon_3)}{\sigma(a_i + \epsilon_2) \times \sigma(a_i + \epsilon_3)} = \frac{1}{\sqrt{\frac{1}{k_2^2}} \times \sqrt{\frac{1}{k_3^2}}} = k_2 \times k_3$ .

I simulate the joint ability and grade distributions 1000 times to get three sets of posterior distributions:

- $f(a_i|g_{i2}, SES)$ , plotted in Figure 1.A.6
- $f(a_i|g_{i3}, SES)$ , plotted in Figure 1.A.7
- $f(a_i|g_{i3}, g_{i2}, SES)$ , plotted in Figures 1.A.8 and 1.A.9

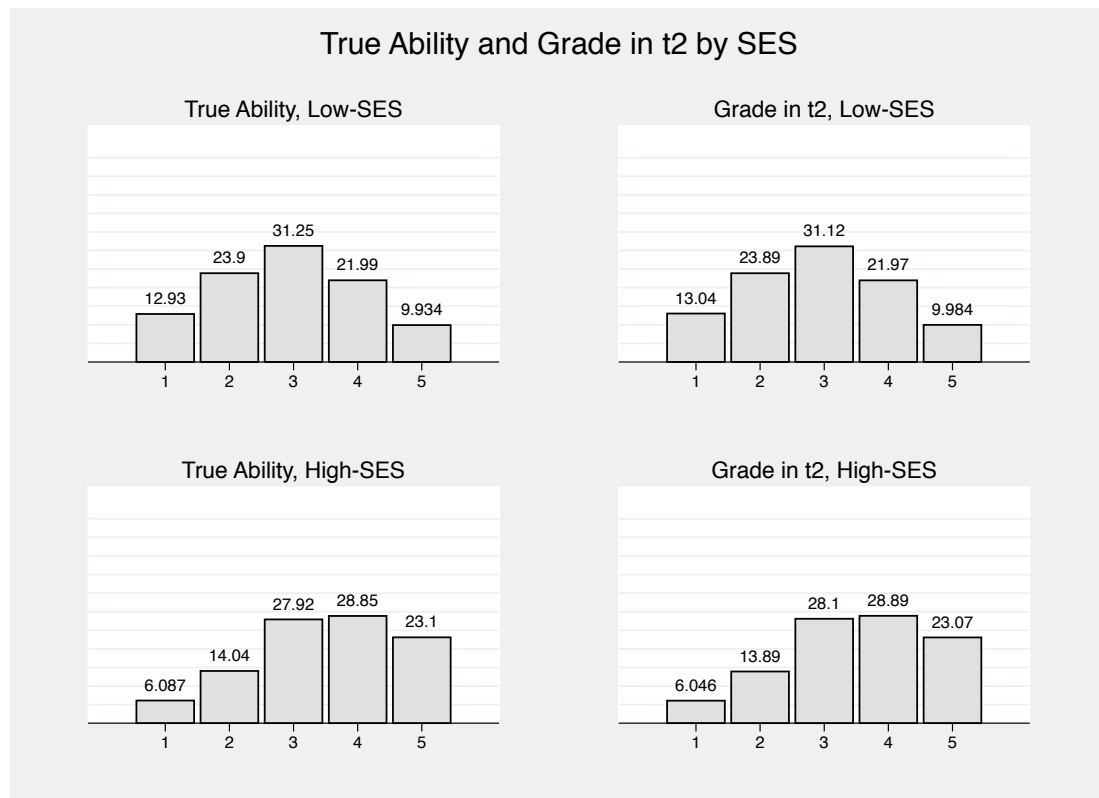


Figure 1.A.5: Ability and grade signals in  $t_2$

Note: The Figure plots simulated distributions of ability and grades in middle compulsory school, for low- and high-SES students.

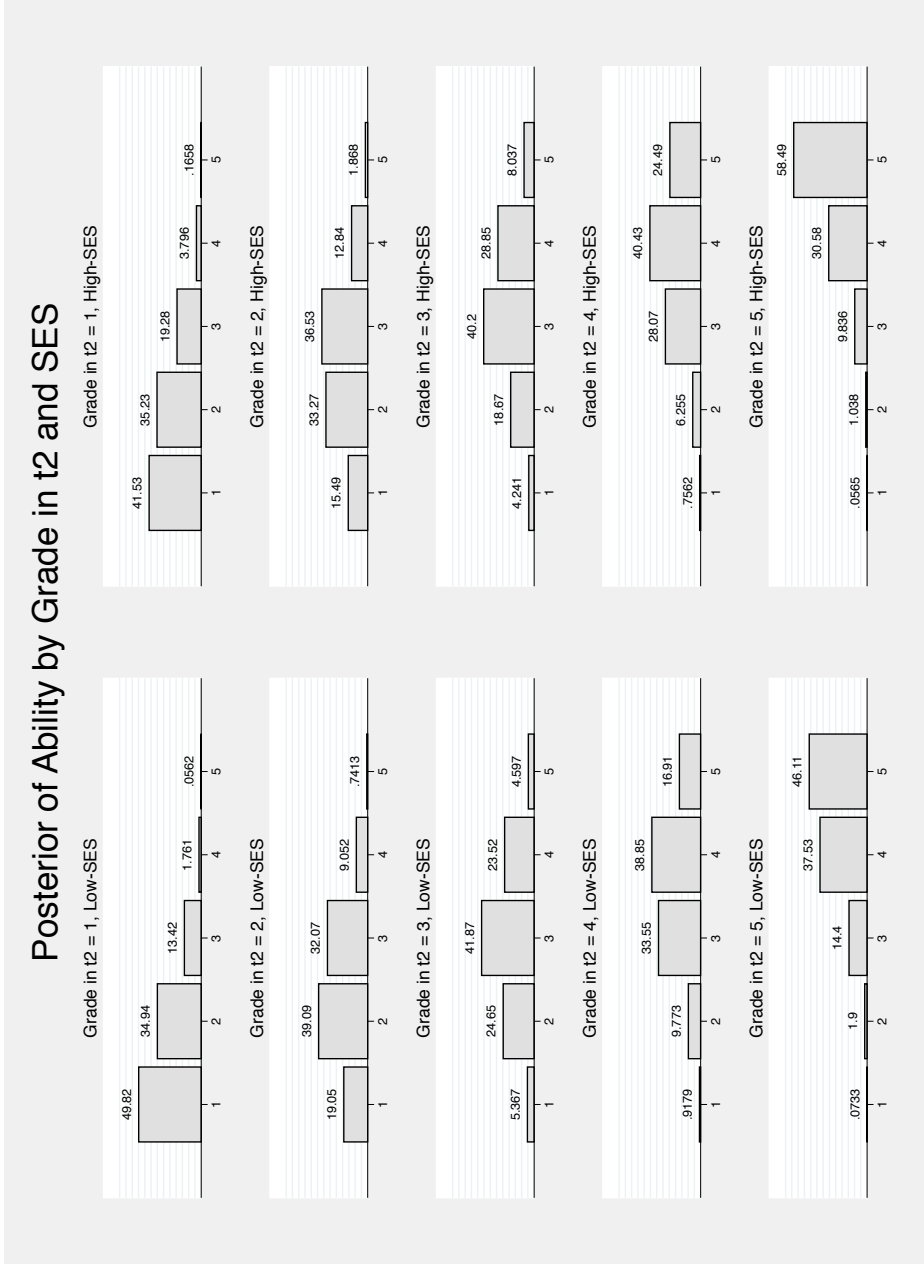


Figure 1.A.6: Posterior distributions by grade signal in  $t_2$

Note: The Figure plots the updated distributions of ability after students observe specific grades in middle compulsory school. Updates differ by SES due to the different priors about ability.

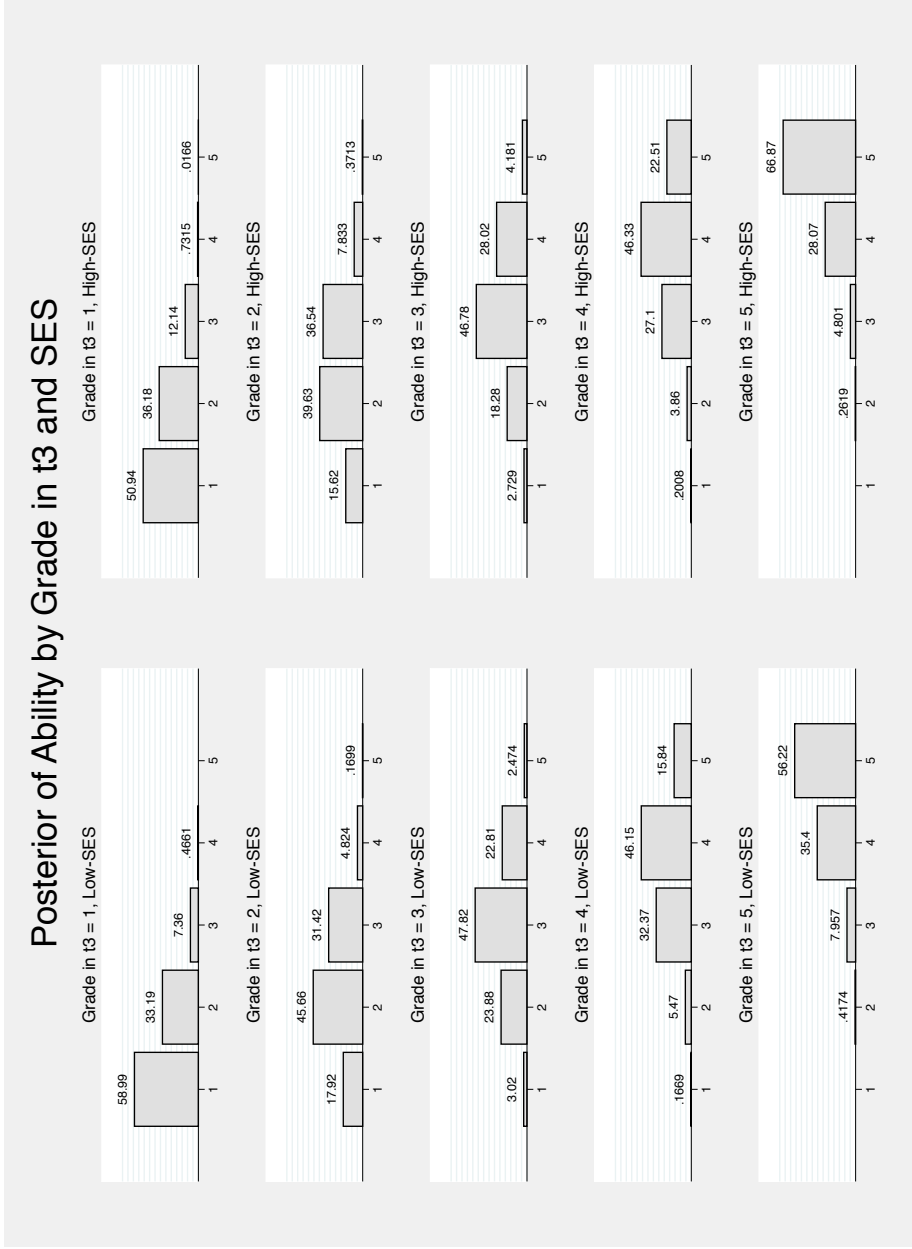
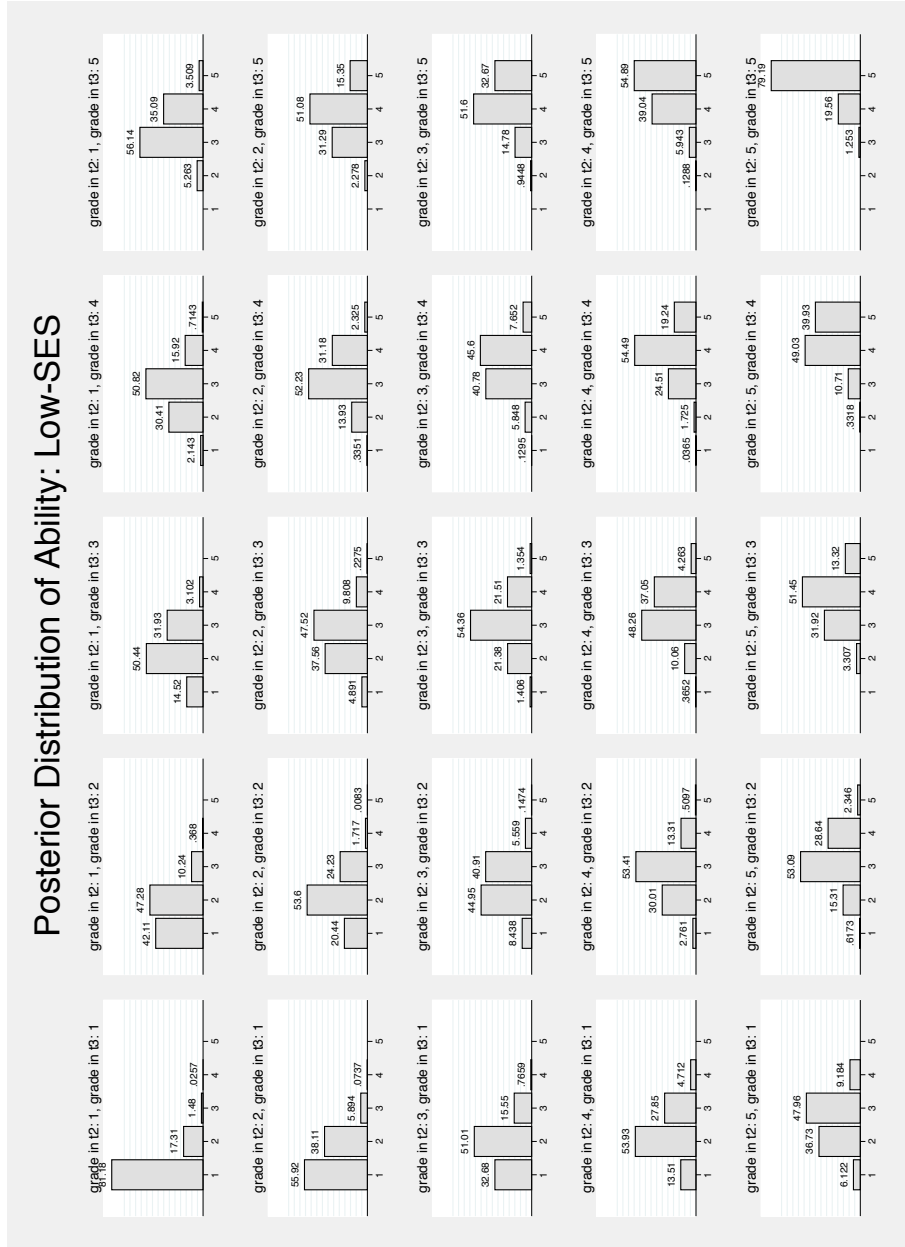


Figure 1.A.7: Posterior distributions by grade signal in  $t_3$

Note: The Figure plots the updated distributions of ability after students observe specific grades in late compulsory school, when only late grades are assigned. Updates differ by SES due to the different priors about ability.



Figure 1.A.8: Posterior distributions after grades in  $t_2$  and  $t_3$  are assigned: Low-SES students

Note: The Figure plots the updated distributions of ability after low-SES students observe specific grades in middle and late compulsory school. Updates differ by SES due to the different priors about ability.

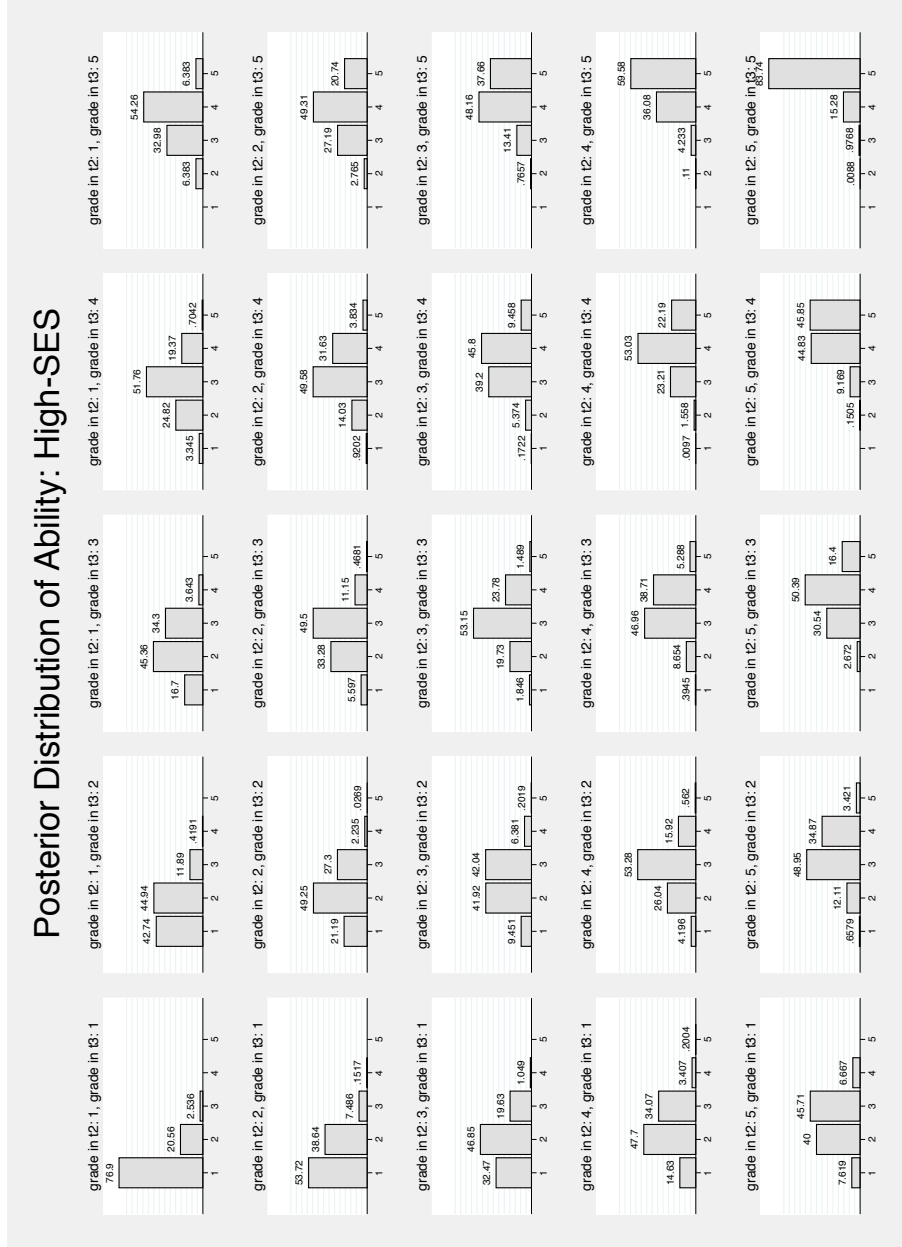


Figure 1.A.9: Posterior distributions after grades in  $t_2$  and  $t_3$  are assigned: High-SES students

Note: The Figure plots the updated distributions of ability after high-SES students observe specific grades in middle and late compulsory school. Updates differ by SES due to the different priors about ability.

### Solution strategy

I solve by backward induction the optimization problem in three different information scenarios:

1. In the case of full information about ability (first best). The solution is found for 5 ability levels. SES has no role in individual choice, but aggregate outcomes will differ due to the different distribution of ability by SES.
2. When only late grades are released. The solution is found for 2 (SES) x 5 ( $g_{i3}$ ) = 10 cases.
3. When early grades are released. The solution is found for 2 (SES) x 5 ( $g_{i2}$ ) x 5 ( $g_{i3}$ ) = 50 cases

### Solution when early grades are assigned

At the end of  $t_3$  choose optimally  $E^*$ , given any  $[SES, e_{i1}, g_{i2}, e_{i2}, g_{i3}, e_{i3}]$  vector. There are  $(2 \times 3 \times 5 \times 3 \times 5 \times 3) \times 3 = 4050$  cases. 1350 solutions are optimal, given  $f(a_i|g_{i3}, g_{i2}, SES)$ . In the same stage choose optimally  $e_{i3}$ , given any  $[SES, e_{i1}, g_{i2}, e_{i2}, g_{i3}]$  vector. There are  $(2 \times 3 \times 5 \times 3 \times 5) \times 3 = 1350$  cases. 450 solutions are optimal given  $f(a_i|g_{i3}, g_{i2}, SES)$ .

In  $t_2$  choose optimally  $e_{i2}$ , given any realized  $[SES, e_{i1}, g_{i2}]$  vector. Use  $f(a_i|g_{i2}, SES)$  to assign the proper weight to each of the 5 potential grades that can be assigned in  $t_3$ .  $E_{max2}$  thus summarizes 450 cases into  $450/5=90$  cases, before  $g_{i2}$  is assigned. There are indeed  $(2 \times 3 \times 5) \times 3 = 90$  cases. 30 solutions are optimal, given  $f(a_i|g_{i2}, SES)$ .

In  $t_1$  choose optimally  $e_1$ , given  $[SES]$ . Use  $f_1(a_i)$  to assign the proper weight to each of the 5 potential grades that might be assigned in  $t_2$ . Thus  $E_{max1}$  summarizes 30 cases into  $30/5=6$  cases, before grades are assigned. There are indeed  $(2) \times 3 = 6$  cases. 2 solutions are optimal, given  $f_1(a_i)$ , one for each SES. So in the end I find  $2 \times 25 [e_1^*, e_2^*, e_3^*, E^*]$  contingent plans, for each SES and ability signal realized.

### Solution when late grades are assigned

At the end of  $t_3$  choose optimally  $E^*$ , given any  $[SES, e_{i1}, e_{i2}, g_{i3}, e_{i3}]$  vector. There are  $(2 \times 3 \times 3 \times 5 \times 3) \times 3 = 810$  cases. 270 cases are optimal, given  $f(a_i|g_{i3}, SES)$ . In  $t_3$  choose optimally  $e_{i3}$ , given any  $[SES, e_{i1}, e_{i2}, g_{i2}]$  vector.

There are  $(2 \times 3 \times 3 \times 5) \times 3 = 270$  cases. 90 solutions are optimal given  $f(a_i|g_{i3}, SES)$ .

In  $t_1$  and  $t_2$  choose optimally  $[e_{i1}, e_{i2}]$ , given  $SES$ . Use  $f_1(a_i)$  to assign the proper weight to each of the 5 potential grades (mirroring ability type) that can be assigned in  $t_3$ . Thus  $Emax_1$  summarizes 90 cases into  $90/5=18$  cases, before grades are assigned. There are indeed  $(2 \times 3) \times 3 = 18$  cases. 2 solutions are optimal given  $f_1(a_i)$ , one for each SES. So in the end I find  $2 \times 5 [e_1^*, e_2^*, e_3^*, E^*]$  contingent plans, for each SES and ability signal realized.

## Realizations

I append the datasets created in the simulation phase, and take a random sample. I merge the final dataset to first and second best solutions. The merge is on  $[SES, g_{i2}, g_{i3}]$  for the solution with early grades,  $[SES, g_{i3}]$  for the solution with late grades, and  $[a_i]$  for the first best. I use true ability, the knowledge production function, and education thresholds, to determine final outcomes. This gives me a distribution of realized outcomes for each SES and ability level. At this point I can assess how the information structure affects final outcomes.

### 1.A.3 Additional Simulation Results

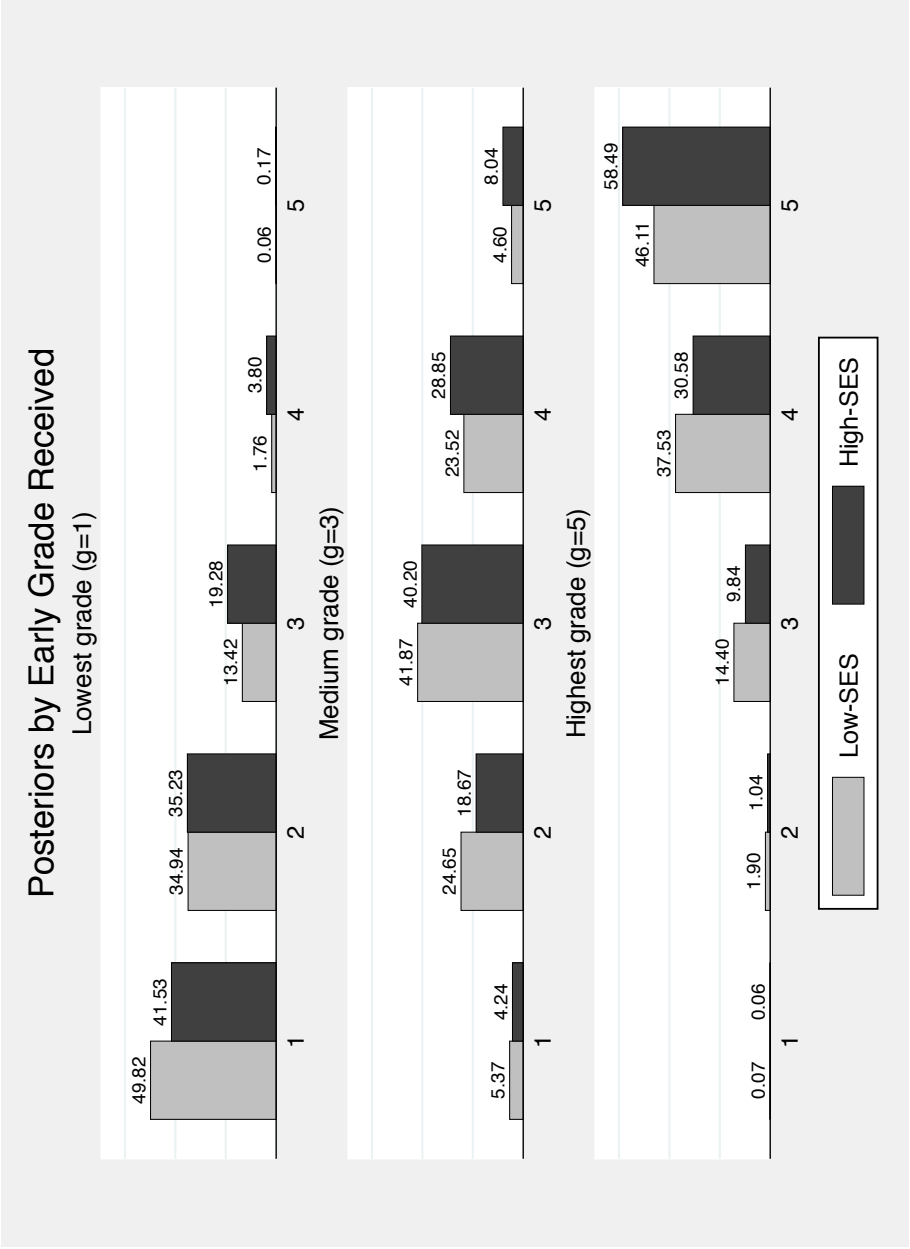


Figure 1.A.10: Comparison of posterior distributions in  $t_2$  after observing grades for low- and high-SES students

Note: The Figure compares posterior distributions for low- and high-SES students in  $t_2$ , after students observe respectively lowest, medium, and top grades. Lowest (highest) grades lead to stronger updates for low (high) SES students.

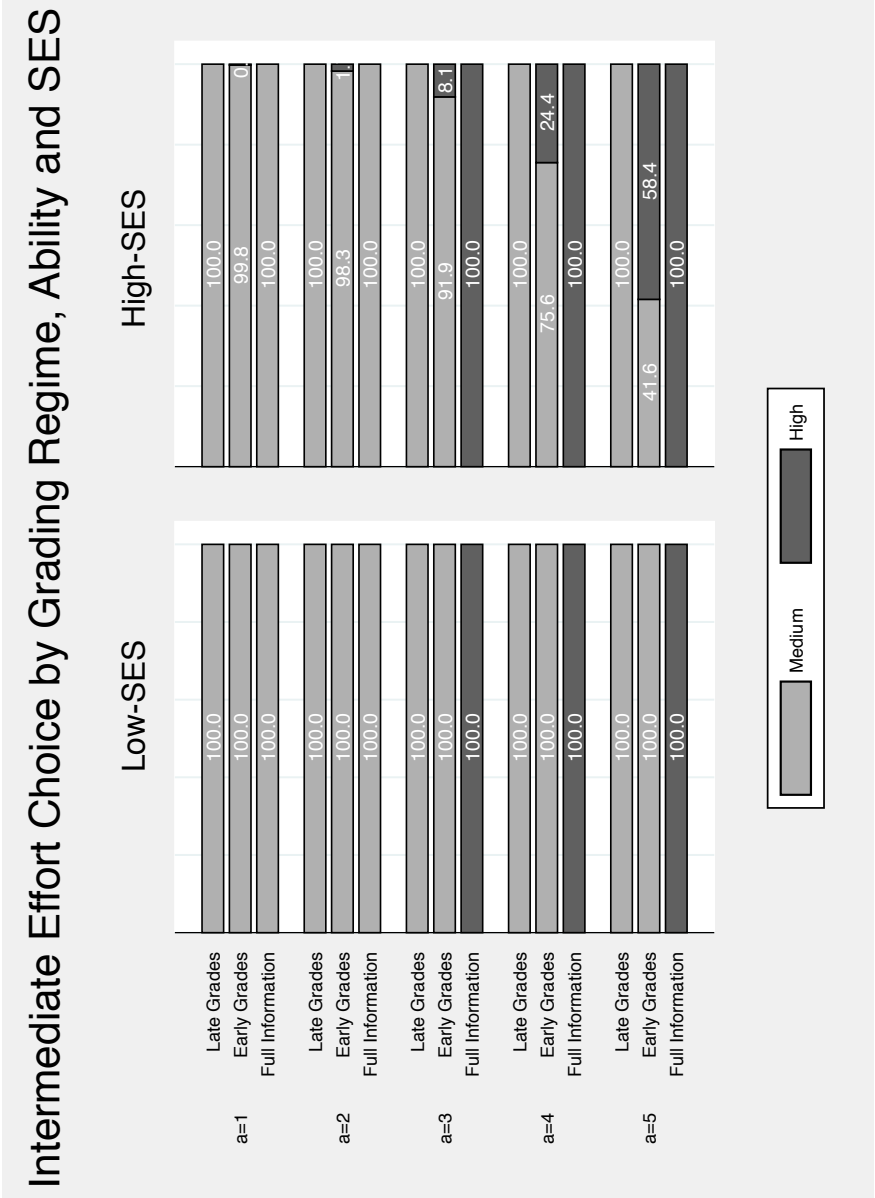


Figure 1.A.1.1: Intermediate effort choice by ability and SES for different grading regimes

Note: The Figure plots effort distributions in middle compulsory school. Results are presented by ability level. For students with ability level above 3, it is optimal to follow academic paths. For those with lower ability, vocational high school is the optimal choice. SES affects students's priors about ability, and thus optimal choices.

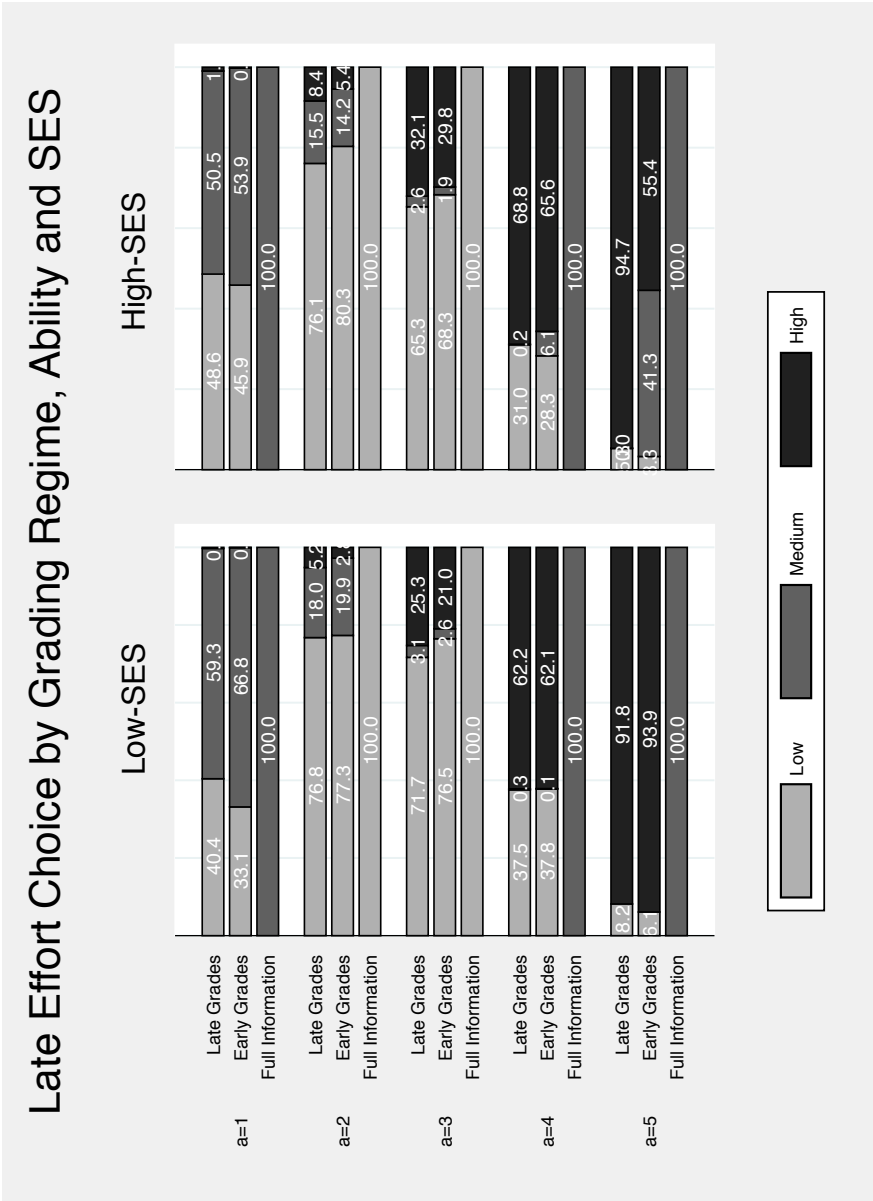


Figure 1.A.12: Late effort choice by ability and SES for different grading regimes

Note: The Figure plots effort distributions in late compulsory school. Results are presented by ability level. For students with ability level above 3, it is optimal to follow academic paths. For those with lower ability, vocational high school is the optimal choice. SES affects students's priors about ability, and thus optimal choices.

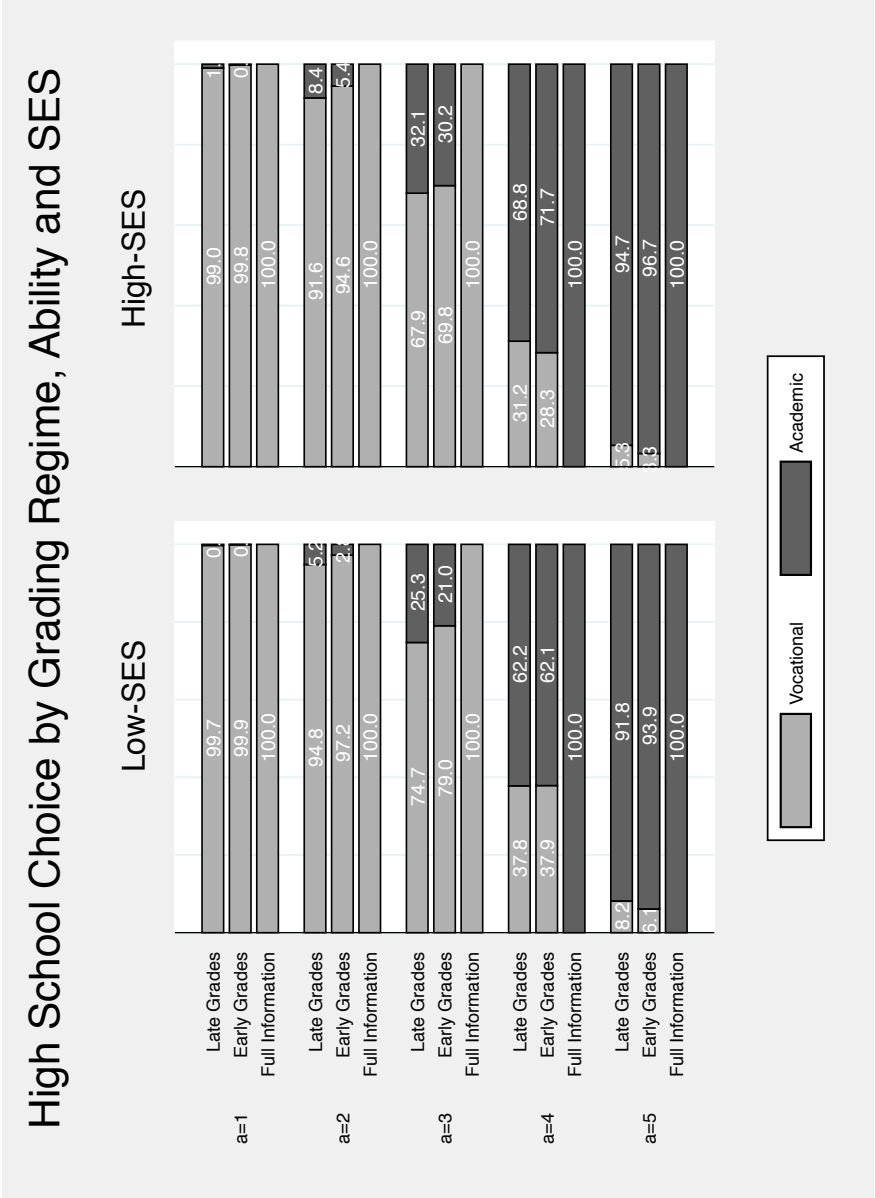


Figure 1.A.13: High school choice by ability and SES for different grading regimes

Note: The Figure plots high school choice distributions. Results are presented by ability level. For students with ability level above 3, it is optimal to follow academic paths. For those with lower ability, vocational high school is the optimal choice. SES affects students's priors about ability, and thus optimal choices.



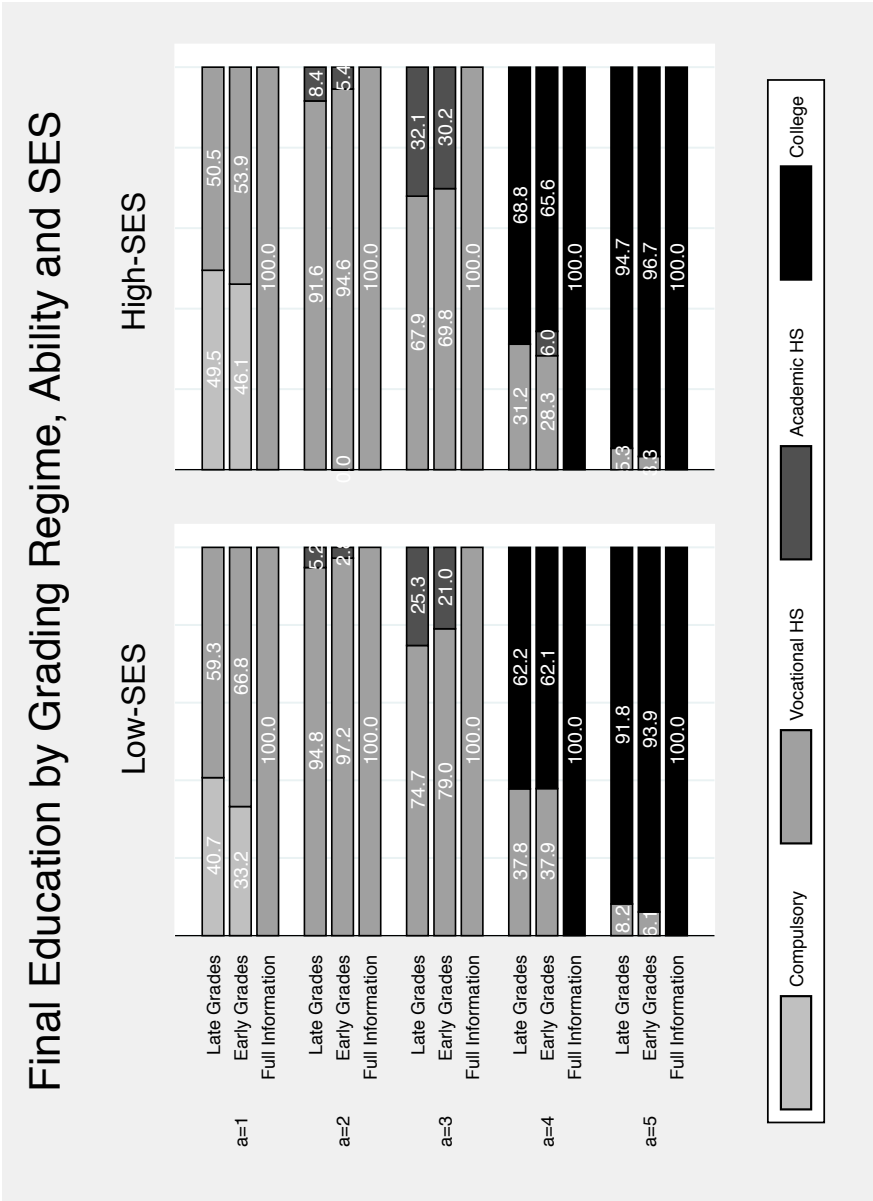


Figure 1.A.14: Final education by ability and SES for different grading regimes

Note: The Figure plots final education distributions. Results are presented by ability level. For students with ability level above 3, it is optimal to follow academic paths. For those with lower ability, vocational high school is the optimal choice. SES affects students's priors about ability, and thus optimal choices.

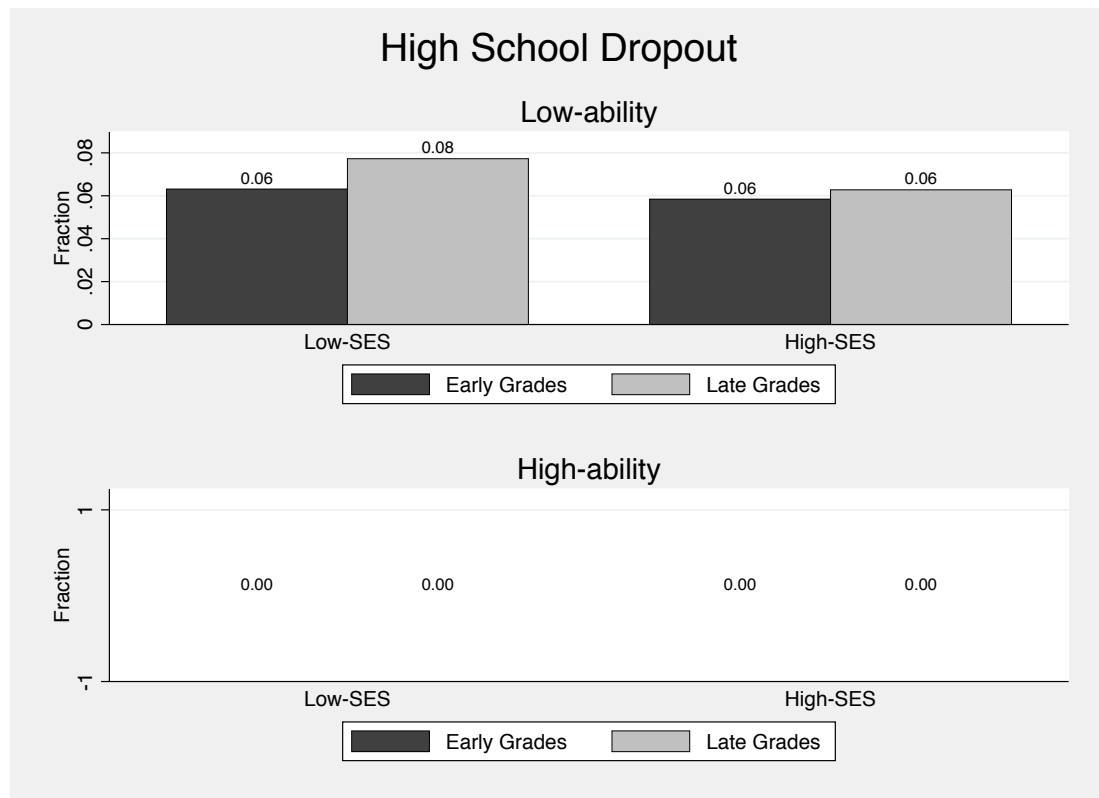


Figure 1.A.15: High school dropout by grading regime

Note: The Figure plots high school dropout rates under early or late grade assignment. Notice that with full information dropout is never optimal. High-ability students never drop out of high school, due to their high levels of ability.

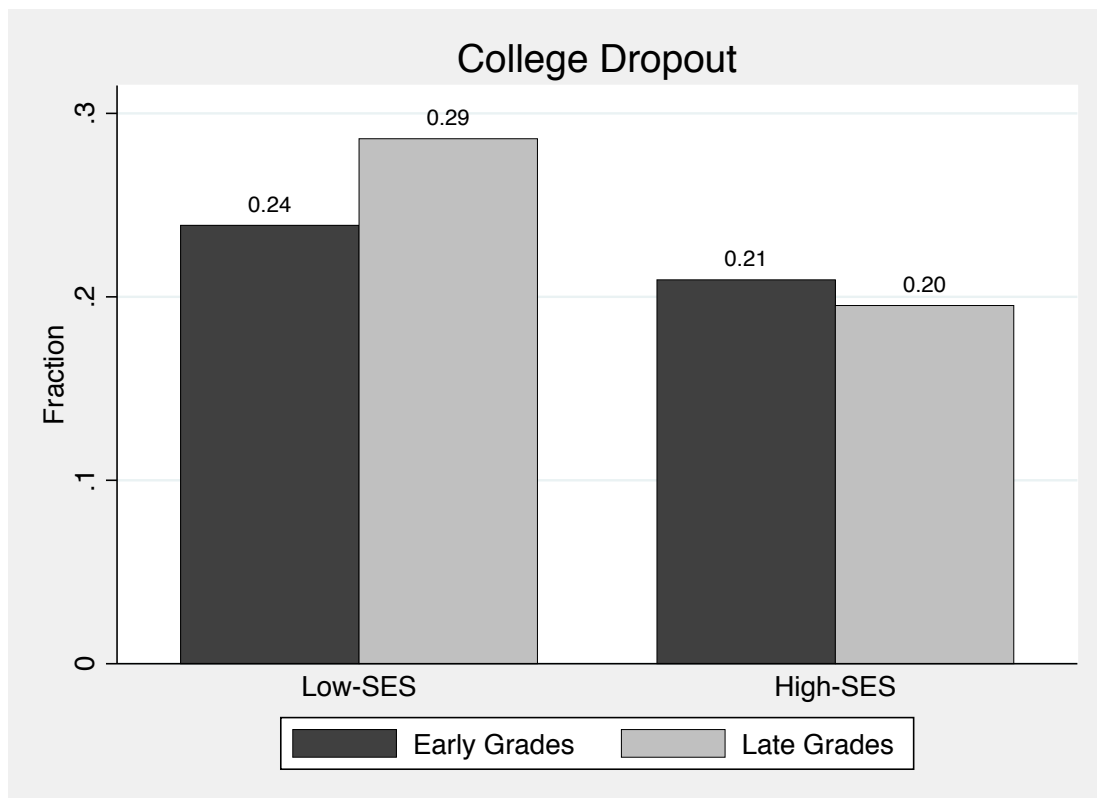


Figure 1.A.16: College dropout by grading regime

Note: The Figure plots college dropout rates under early or late grade assignment. Notice that with full information dropout is never optimal.

### 1.A.4 Model and Institutional Setup

Table 1.A.7 compares the model to the institutional setup in terms of choices, selection and information. While I designed the model around the institutional setup, there still are some differences.

Table 1.A.7: Model and empirical setup

	Model	Empirical setup
<b>Early Compulsory</b>	choose $e_1$ , no grades	choose effort in s.y. 1-5, no grades
<b>Middle Compulsory</b>	choose $e_2$ , (no) grades	choose effort in s.y. 6 (no) grades
<b>Late Compulsory</b>	choose $e_3$ , grades	choose effort/courses in s.y. 7-9, grades
<b>High school Selection</b>	$\bar{k}^{E_3}$	GPA, course choices
<b>College Selection</b>	$E_3, \bar{k}^{E_4}$	Academic HS, GPA, quotas

First, students in the sample have an additional choice with respect to the model: type of course in late compulsory school. As advanced courses are more difficult than general ones, this type of choice can be interpreted as an effort choice. At the same these choices are relevant for admission into academic high school. Table 1.A.8 confirms this empirically: choosing advanced math in grade 9 and having higher grades substantially affects the probability of admission to the preferred high school choice. The magnitudes of the coefficients are lower bounds, as I am including students who only apply to vocational tracks.<sup>39</sup>

Second, in the model I assume that students need to meet absolute knowledge thresholds to complete college. In my setup a quota system is in place: an increase in college enrollment could in principle affect the admission threshold for all students. These general equilibrium effects are not captured in my model, where the number of students who can complete college is a function of the ability distribution. Öckert (2002) finds that the difference in years of education between students screened out and admitted at college in the early 80s in Sweden is about 0.6. The difference reduces to just 0.20 years when comparing students with similar number of admission credits.<sup>40</sup> This in turn is a good approximation to the marginal change in admission requirements that might be triggered by a reshuffling of the pool seeking college admission after grades are assigned. Given the size of the change results would likely not change

<sup>39</sup>I have no information on the track the student applied for.

<sup>40</sup>See Öckert (2010), published version of the IFAU working paper.

significantly allowing for the general equilibrium effect.

Table 1.A.8: Factors Affecting HS admission,  
Cohort 1967

	Admitted in HS at first Choice
Adv Math (s.y. 9)	0.05*** (0.01)
Adv English (s.y. 9)	0.00 (0.02)
Math Grade (s.y. 9)	0.04*** (0.01)
English Grade (s.y. 9)	0.02*** (0.01)
Swedish Grade (s.y. 9)	0.03*** (0.01)
Mean	0.83
$R^2$	0.05
Observations	7884

Note: \*  $p < 0.10$ , \*\*  $p < 0.05$ , \*\*\*  $p < 0.01$

The data does not record which type of school the student was applying for. Standard errors clustered at the class level.

## 1.B Descriptives

### 1.B.1 Definition of Ability and SES

In this section I discuss how I measure ability and SES, and describe how I discretize them to match the model.

Students took during the spring term of school year 6 a battery of three standardized ability tests: a test of *verbal ability*, requiring to find the opposite

to a word among a list of four alternatives; a test of *inductive ability*, requiring to complete a number series of 6 terms with two more numbers; a test of *spatial ability*, requiring to find the three-dimensional representation of a two-dimensional picture that can be folded. The tests taken by the two cohorts are exactly the same, and the distributions look similar over time (see Figure 1.B.1).

Students had respectively 15, 27 and 22 seconds to answer each section of the test, assuming they wasted no time at all in the test. The fast pace of the test adds to the quality of the ability measures: Borghans et al. (2008) show that reducing the time available for completing intelligence tests reduces differences in effort between students with different non-cognitive traits.

I create a standardized aggregate index of ability from the z-scores of inductive and verbal ability. I label high-ability those students who scored at least at the 60th percentile of the ability distribution. Consistently with the model, the cutoff roughly corresponds to the median ability of students who attained college education.<sup>41</sup>

When performing the normalization at the cohort level, ability measures turn out to be 5% of a standard deviation higher in the treatment group, with respect to the control group. For the 1972 cohort there is no such difference. While the main treatment is grade assignment in school year 6, in principle there might be differences in grade assignment also in school year 3, when students were age 10 (see discussion in Section 1.2.3). Ability could thus have been affected by the treatment. However the literature (e.g., Heckman et al., 2007) reports that cognitive ability should be stable by that age. The ability measures were taken in May of 1981 for the 1967 cohort, quite close to the final tests used for grade assignment. It is possible that test taking behavior was instead affected in the treated municipalities. Students may have thought that the standardized tests were relevant to the final grades, or may have put more effort in the tests simply because they were affected by the more competitive environment.<sup>42</sup> This is consistent with results from the literature on the effect of non-cognitive traits on test taking behavior (Borghans et al., 2008). In order to have more consistent ability measures I thus normalize ability at the cohort-treatment level, and basically use a measure of relative ability in the analysis.

---

<sup>41</sup>I leave out of the index spatial ability, as it poorly correlates with academic choices or outcomes.

<sup>42</sup>Jalava et al. (2015) show that rank-based grading positively affects effort during tests

This avoids any problem of endogeneity or differential reporting caused by grade assignment.

Extensive investigation of which SES measure is most predictive of education choice shows that parental education strongly predicts children's education choices. Parental income is less predictive of education choice. Measures based on parental occupation yield results similar to parental education. My preferred measure of SES is based on parental education. Occupation-based measures are more difficult to discretize into dummies, and are recorded in my data using a definition that slightly changes between cohorts.<sup>43</sup> Finally, I consider high-SES those students who have at least one parent with academic high school (about 40% of my sample).

---

<sup>43</sup>Results do not change that much when using the alternative SES definition

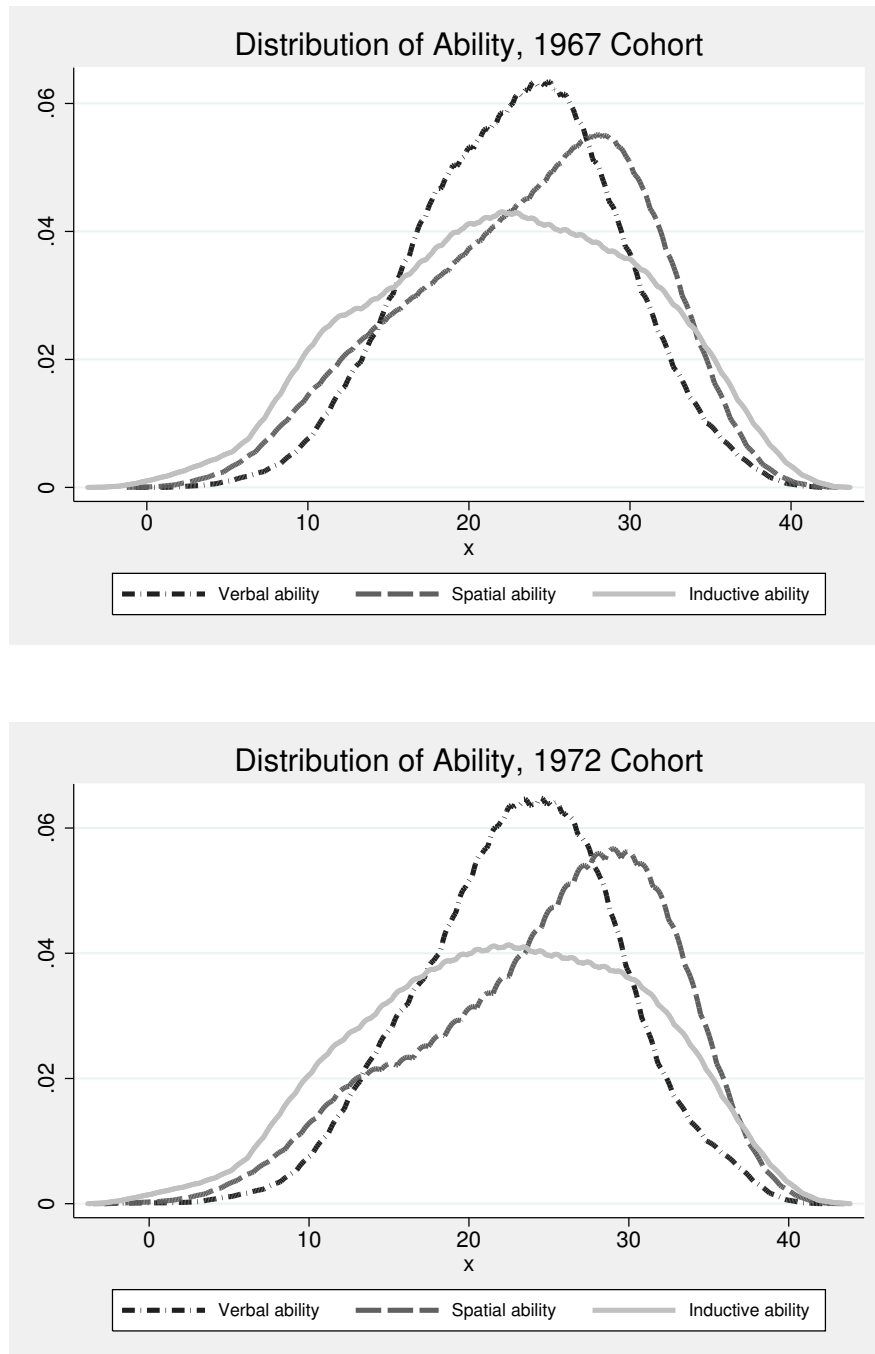


Figure 1.B.1: Absolute ability distributions

Note: Ability measures are taken from tests administered in school year 6.  
The tests are the same for both cohorts.



## 1.B.2 Education Choices, Grades and Outcomes

In the following I provide descriptive evidence on students' choices and educational attainment, the main outcomes in the empirical analysis.

Table 1.B.1: Education choices and outcomes by ability and SES, cohort 1967

	Low-ability		High-ability	
	Low-SES	High-SES	Low-SES	High-SES
<b>Compulsory:</b>				
Adv Math (s.y. 7)	0.56	0.73	0.92	0.97
Adv Math (s.y. 8)	0.44	0.62	0.87	0.95
Adv Math (s.y. 9)	0.32	0.51	0.75	0.88
Adv English (s.y. 7)	0.57	0.77	0.93	0.98
Adv English (s.y. 8)	0.53	0.74	0.90	0.97
Adv English (s.y. 9)	0.47	0.66	0.86	0.95
<b>Non Compulsory:</b>				
HS Enroll	0.85	0.92	0.93	0.98
Academic HS Enroll	0.22	0.45	0.60	0.82
HS Dropout	0.19	0.12	0.10	0.09
Attain Academic HS	0.19	0.39	0.56	0.75
Attain College	0.19	0.39	0.48	0.72

Each variable is a dummy. College enrollment is defined as enrolling into a tertiary education program lasting at least 2 years.

Table 1.B.1 shows that high-ability and high-SES students are more likely to make choices consistent with an academic education path. This pattern suggests that the Swedish education system (grading, tracking and funding) successfully managed to allocate the most skilled students to higher education levels, but SES remained a relevant factor in the process, possibly distorting the efficient allocation of skill to human capital. While these differences are less marked for high-ability students up to compulsory school, low-SES students are 20 p.p. less likely to choose an academic education, independently of ability. As there might still be differences in ability between SES categories, in Table 1.B.2 I

report coefficients for the differences in choices between high- and low-SES students, controlling for ability. The picture does not change that much, but it appears that the differences among low ability students are in part due to low-SES students having less ability. I still confirm that low-SES students are much less likely (a 18 p.p difference) to choose an academic education path than their high-SES counterparts.

Table 1.B.2: Differences in choices by SES  
controlling for ability, cohort 1967

	Low-ability	High-ability
<b>Compulsory:</b>		
Adv Math (s.y. 7)	0.11	0.04
Adv Math (s.y. 8)	0.12	0.07
Adv Math (s.y. 9)	0.13	0.10
Adv English (s.y. 7)	0.14	0.04
Adv English (s.y. 8)	0.15	0.06
Adv English (s.y. 9)	0.14	0.07
<b>High School:</b>		
HS Enroll	0.06	0.04
Academic Enroll	0.19	0.18
HS Dropout	-0.05	-0.00
Attain Academic HS	0.17	0.16
<b>College:</b>		
College	0.17	0.21

Each variable is a dummy. College enrollment is defined as enrolling into a tertiary education program lasting at least 2 years.

To understand whether differences in education choices and educational attainment are related to school performance, I report in Table 1.B.3 average grades by ability and SES. Grades are consistently higher for high-ability and high-SES students. However grade differences between high- and low-SES students with similar ability levels are not so big. When considering grades in school year 6, and Swedish grades in school year 9 (which do not reflect course choice), differences are at most  $\frac{1}{3}$ th of a grade. Table 1.B.4 reports grade

differences by SES controlling for ability. The picture remains similar: grade differences among students with similar ability levels are at most  $\frac{1}{4}$ th of a grade. This suggests that SES plays a fundamental role in education choice in Sweden, potentially reflecting different motivation and preferences for education among students.

Table 1.B.3: Grades by ability and SES, cohort 1967

	Low-ability		High-ability	
	Low-SES	High-SES	Low-SES	High-SES
Swedish Grade (s.y. 6)	2.64	2.93	3.67	3.97
Swedish Grade (s.y. 7)	2.61	2.86	3.40	3.63
Swedish Grade (s.y. 8)	2.64	2.88	3.46	3.72
Swedish Grade (s.y. 9)	2.69	3.01	3.56	3.89
Math Grade (s.y. 6)	2.66	2.97	3.80	3.98
Math Grade (s.y. 7)	2.64	2.78	3.38	3.61
Math Grade (s.y. 8)	2.69	2.86	3.37	3.61
Math Grade (s.y. 9)	2.83	3.00	3.51	3.72
English Grade (s.y. 6)	2.64	2.98	3.65	4.02
English Grade (s.y. 7)	2.69	2.84	3.33	3.63
English Grade (s.y. 8)	2.68	2.86	3.37	3.63
English Grade (s.y. 9)	2.80	3.06	3.45	3.76

Grades are expressed on a 1-5 norm-referenced scale. Math and English grades in s.y. 8 and 9 pool together advanced and general courses.

Table 1.B.5 shows that many students switch courses over time. Students are more likely to switch from academic to general courses than the opposite, and there are more switches in math.<sup>44</sup> Switches from one type of course to the other can be interpreted as revision of choice, and imply that students do not have full information over own ability (or knowledge). It is also interesting to see that the switching behavior continues through all grades: most of the students switch in school year 8, but some also switch in the last year. This is consistent

<sup>44</sup>This is consistent with research finding that students tend to be overly optimistic about own ability/preparation in higher education (see Stinebrickner & Stinebrickner, 2012; Zafar, 2011)

Table 1.B.4: Differences in grades by SES  
controlling for ability, cohort 1967

	Low-ability	High-ability
Swedish Grade (s.y. 6)	0.18	0.21
Swedish Grade (s.y. 7)	0.16	0.15
Swedish Grade (s.y. 8)	0.15	0.17
Swedish Grade (s.y. 9)	0.22	0.24
Math Grade (s.y. 6)	0.18	0.11
Math Grade (s.y. 7)	0.07	0.14
Math Grade (s.y. 8)	0.11	0.16
Math Grade (s.y. 9)	0.10	0.14
English Grade (s.y. 6)	0.22	0.27
English Grade (s.y. 7)	0.07	0.21
English Grade (s.y. 8)	0.11	0.16
English Grade (s.y. 9)	0.20	0.21

Grades are expressed on a 1-5 norm-referenced scale. Math and English grades in s.y. 8 and 9 pool together advanced and general courses.

with students revising some sort of prior, with the updating process continuing over time. In Table 1.B.4 I show how SES affects the choice of switching in compulsory school from an advanced to a general course. When comparing the choices of low- and high-SES students with the same grades and ability, I find that SES still influences switching choices. This suggests that grading information might affect differently students from different socioeconomic backgrounds.

Table 1.B.5: Fraction of students switching courses, 1967 cohort

	Grade 8	Grade 9
<b>Math:</b>		
Switches to gen choice	0.12	0.17
Switches to adv choice	0.04	0.01
<b>English:</b>		
Switches to gen choice	0.06	0.08
Switches to adv choice	0.07	0.03

Switches are conditional on previous year's course choice.

Table 1.B.6: Impact of SES on the stability of course choices, 1967 Cohort

	Sticks to adv Math	Sticks to adv Eng
<b>Regressors of interest:</b>		
High-SES	0.06*** (0.01)	0.02* (0.01)
<b>Controls:</b>		
Grade (s.y. 8)	-0.00 (0.01)	-0.05*** (0.01)
Grade (s.y. 7)	0.21*** (0.01)	0.18*** (0.01)
Standardized verbal ability	0.02** (0.01)	0.04*** (0.01)
Standardized inductive ability	0.02** (0.01)	0.00 (0.01)
Standardized spatial ability	0.04*** (0.01)	0.02*** (0.01)
E[Y]	0.73	0.86
R <sup>2</sup>	0.27	0.21
Observations	5532	5653

Note: \* p < 0.10, \*\* p < 0.05, \*\*\* p < 0.01

Outcome: the student sticks to the advanced course choice made in s.y. 7. The SES division is based on parental education. Standard errors clustered at the class level.

## 1.B.3 Treated and Control Municipalities

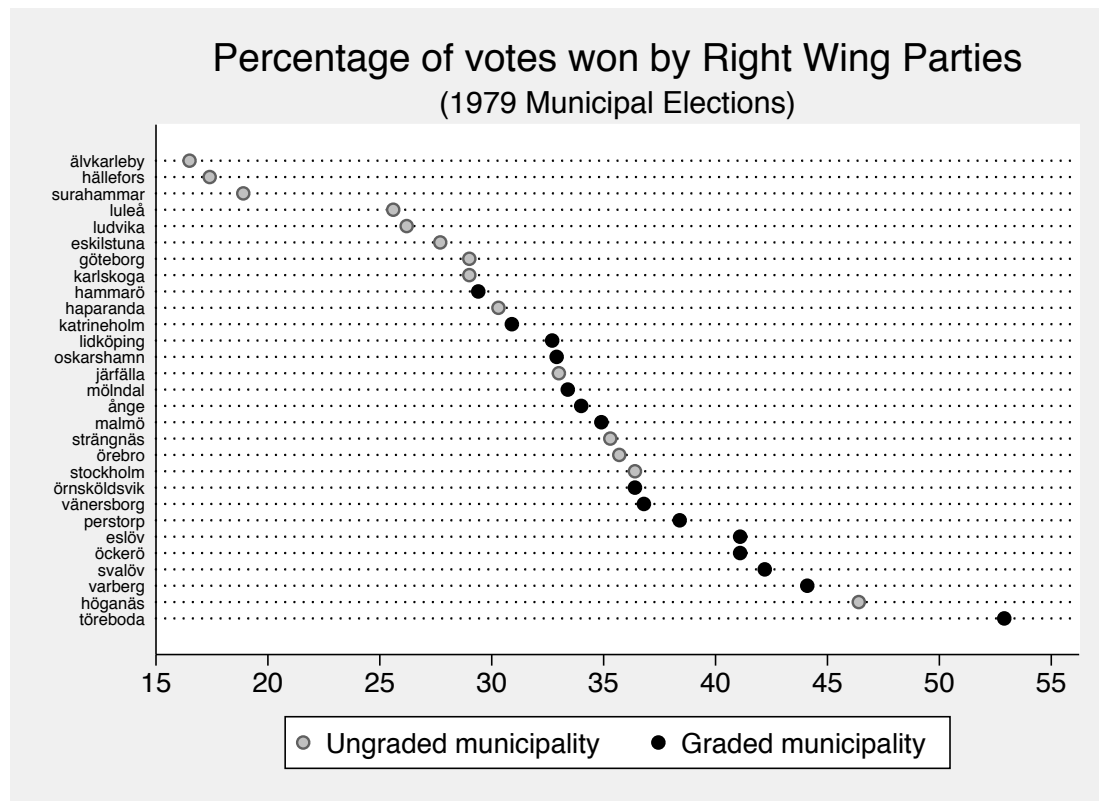


Figure 1.B.2: Vote share of right-wing parties in municipal elections by treatment status

Note: This Figure plots the aggregated vote share of right-wing parties in the 1979 municipal elections. Municipalities assigning early grades had a higher share of right-wing voters.

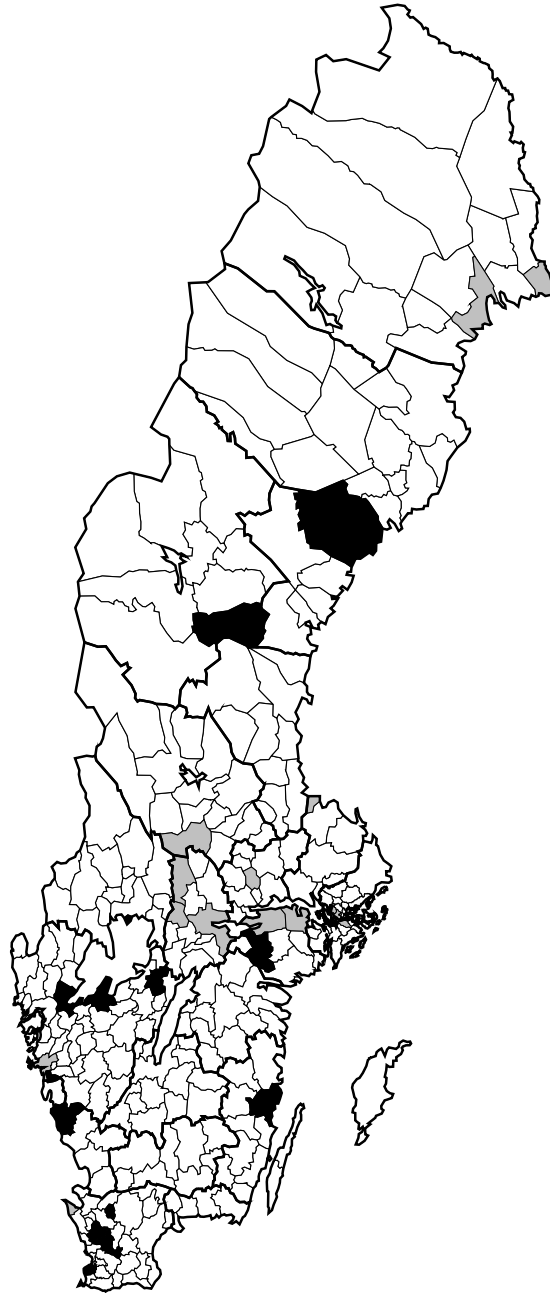


Figure 1.B.3: Sampled municipalities

Note: Municipalities assigning early grades before the reform are colored in black

Table 1.B.7: Differences in students background by treatment status, 1967 cohort

	Graded	Ungraded	Difference
Female	0.48 (0.62)	0.50 (0.55)	-0.02 (0.01)
Birth year	66.98 (0.25)	66.97 (0.24)	0.02* (0.01)
Foreign born	0.02 (0.14)	0.05 (0.24)	-0.03*** (0.01)
Both parents not Nordic	0.01 (0.11)	0.02 (0.13)	-0.00 (0.01)
Verbal ability	23.22 (7.60)	22.70 (6.60)	0.52* (0.28)
Inductive ability	22.85 (10.25)	21.78 (8.92)	1.07*** (0.38)
Spatial ability	23.82 (8.94)	23.41 (7.91)	0.41 (0.28)
Kindergarten	0.91 (0.34)	0.93 (0.28)	-0.02 (0.03)
Quiet home environment	0.95 (0.28)	0.95 (0.25)	-0.00 (0.01)
Switched Class (G6-G9)	0.06 (0.27)	0.08 (0.30)	-0.03*** (0.01)
Special Education	0.12 (0.41)	0.14 (0.39)	-0.02 (0.02)
Changes of teacher	0.59 (1.20)	0.53 (1.00)	0.05 (0.14)
Hours absent in grade 6	7.89 (11.94)	6.82 (8.23)	1.07 (1.26)
Class size	23.64 (6.64)	23.27 (5.18)	0.37 (0.75)

\*  $p < 0.10$ , \*\*  $p < 0.05$ , \*\*\*  $p < 0.01$

Standard errors clustered at the municipal level. Ability measures are on a 0-40 scale.



Table 1.B.8: Differences in parental SES by treatment status, 1967 cohort

	Graded	Ungraded	Difference
Disposable family income (age 11-16)	243.69 (101.46)	253.82 (84.98)	-10.13** (4.87)
High Income	0.47 (0.62)	0.54 (0.55)	-0.07** (0.03)
High Education	0.36 (0.60)	0.44 (0.54)	-0.08** (0.03)
High Income/Educ	0.59 (0.62)	0.68 (0.51)	-0.09*** (0.03)
High SES	0.41 (0.61)	0.46 (0.55)	-0.05 (0.04)
Parents: non-skilled workers, goods	0.10 (0.39)	0.09 (0.33)	0.01 (0.02)
Parents: non-skilled workers, service	0.11 (0.39)	0.09 (0.31)	0.02** (0.01)
Parents: skilled workers, goods	0.17 (0.47)	0.19 (0.43)	-0.02 (0.02)
Parents: skilled workers, service	0.02 (0.18)	0.02 (0.16)	-0.00 (0.01)
Parents: lower non-manual ii	0.04 (0.25)	0.05 (0.22)	-0.00 (0.01)
Parents: lower non-manual i	0.09 (0.35)	0.09 (0.30)	0.00 (0.01)
Parents: intermediate-level non-manual	0.19 (0.49)	0.23 (0.46)	-0.03** (0.01)
Parents: higher civil servants and senior salaried	0.11 (0.39)	0.13 (0.36)	-0.01 (0.02)
Parents: independent professionals	0.00 (0.03)	0.00 (0.06)	-0.00** (0.00)
Parents: entrepreneur	0.11 (0.39)	0.10 (0.33)	0.01 (0.01)
Parents: farmer	0.05 (0.30)	0.02 (0.16)	0.04*** (0.01)

\*  $p < 0.10$ , \*\*  $p < 0.05$ , \*\*\*  $p < 0.01$

Standard errors clustered at the municipal level. Income in 1000 kr, measured when the student was 11-16. Occupation variables are taken from the 1980 Census.

Table 1.B.9: Differences in father background  
by treatment status, 1967 cohort

	Graded	Ungraded	Difference
Father not Nordic	0.03 (0.19)	0.03 (0.20)	-0.01 (0.01)
Married father	0.81 (0.49)	0.77 (0.46)	0.04*** (0.01)
Father SES, 1 (low) to 3 (high)	1.73 (0.87)	1.75 (0.79)	-0.02 (0.05)
Divorced father	0.15 (0.45)	0.18 (0.42)	-0.03** (0.01)
Father educ: compulsory school or less	0.47 (0.62)	0.41 (0.54)	0.06** (0.03)
Father educ: high school	0.38 (0.60)	0.41 (0.54)	-0.04** (0.02)
Father educ: college or more	0.16 (0.45)	0.18 (0.42)	-0.02 (0.02)
Father: in the labor force	0.90 (0.36)	0.90 (0.33)	0.00 (0.01)
Father: unemployed	0.04 (0.24)	0.04 (0.20)	0.00 (0.00)

\*  $p < 0.10$ , \*\*  $p < 0.05$ , \*\*\*  $p < 0.01$

Standard errors clustered at the municipal level.

Table 1.B.10: Differences in mother background  
by treatment status, 1967 cohort

	Graded	Ungraded	Difference
Mother not Nordic	0.04 (0.21)	0.05 (0.23)	-0.01 (0.01)
Married mother	0.77 (0.52)	0.72 (0.49)	0.04*** (0.02)
Divorced mother	0.16 (0.45)	0.19 (0.43)	-0.04** (0.01)
Mother SES, 1 (low) to 3 (high)	1.55 (0.76)	1.62 (0.68)	-0.06 (0.05)
Mother educ: compulsory school or less	0.42 (0.61)	0.41 (0.54)	0.01 (0.03)
Mother educ: high school	0.41 (0.62)	0.41 (0.54)	0.00 (0.02)
Mother educ: college or more	0.17 (0.47)	0.18 (0.42)	-0.01 (0.02)
Mother: in the labor force	0.91 (0.36)	0.90 (0.34)	0.01 (0.01)
Mother: unemployed	0.04 (0.25)	0.03 (0.19)	0.01* (0.00)

\*  $p < 0.10$ , \*\*  $p < 0.05$ , \*\*\*  $p < 0.01$ 

Standard errors clustered at the municipal level.

Table 1.B.11: Differences in father occupation  
by treatment status, 1967 cohort

	Graded	Ungraded	Difference
Father: non-skilled workers, goods	0.12 (0.42)	0.12 (0.35)	0.01 (0.02)
Father: non-skilled workers, service	0.10 (0.38)	0.08 (0.29)	0.03*** (0.01)
Father: skilled workers, goods	0.21 (0.50)	0.23 (0.46)	-0.02 (0.02)
Father: skilled workers, service	0.01 (0.09)	0.01 (0.10)	-0.00 (0.00)
Father: lower non-manual ii	0.03 (0.21)	0.02 (0.17)	0.01 (0.00)
Father: lower non-manual i	0.07 (0.32)	0.07 (0.29)	0.00 (0.01)
Father: intermediate-level non-manual	0.18 (0.47)	0.22 (0.45)	-0.04** (0.02)
Father: higher civil servants and senior salaried	0.11 (0.38)	0.12 (0.36)	-0.02 (0.02)
Father: independent professionals	0.00 (0.03)	0.00 (0.05)	-0.00* (0.00)
Father: entrepreneur	0.11 (0.39)	0.10 (0.32)	0.01 (0.01)
Father: farmer	0.06 (0.30)	0.02 (0.17)	0.03** (0.01)

\*  $p < 0.10$ , \*\*  $p < 0.05$ , \*\*\*  $p < 0.01$

Standard errors clustered at the municipal level. Occupation variables are taken from the 1980 Census.

Table 1.B.12: Differences in mother occupation  
by treatment status, 1967 cohort

	Graded	Ungraded	Difference
Mother: non-skilled workers, goods	0.05 (0.27)	0.06 (0.26)	-0.00 (0.01)
Mother: non-skilled workers, service	0.37 (0.60)	0.36 (0.53)	0.02 (0.03)
Mother: skilled workers, goods	0.01 (0.14)	0.02 (0.15)	-0.01 (0.01)
Mother: skilled workers, service	0.06 (0.31)	0.06 (0.25)	0.00 (0.01)
Mother: lower non-manual ii	0.11 (0.39)	0.14 (0.38)	-0.03* (0.01)
Mother: lower non-manual i	0.09 (0.34)	0.11 (0.33)	-0.02 (0.01)
Mother: intermediate-level non-manual	0.16 (0.47)	0.16 (0.40)	-0.00 (0.01)
Mother: higher civil servants and senior salaried	0.04 (0.25)	0.04 (0.21)	0.00 (0.01)
Mother: independent professionals	0.00 (0.00)	0.00 (0.03)	-0.00* (0.00)
Mother: entrepreneur	0.05 (0.28)	0.04 (0.23)	0.00 (0.01)
Mother: farmer	0.05 (0.28)	0.02 (0.15)	0.03*** (0.01)

\*  $p < 0.10$ , \*\*  $p < 0.05$ , \*\*\*  $p < 0.01$ 

Standard errors clustered at the municipal level. Occupation variables are taken from the 1980 Census.

Table 1.B.14: Differences in parental school preferences  
by treatment status, 1967 cohort

	Graded	Ungraded	Difference
School priority: Teach Math and Swedish	8.21 (1.73)	8.06 (1.63)	0.14** (0.06)
School priority: Develop collaborative skills	6.64 (2.26)	6.70 (2.02)	-0.06 (0.09)
School priority: Teach other subjects	6.03 (2.35)	6.01 (1.99)	0.02 (0.07)
School priority: Develop critical thinking	5.95 (2.62)	6.17 (2.33)	-0.22** (0.10)
School priority: Teach Foreign languages	5.46 (2.57)	5.61 (2.19)	-0.16 (0.14)
School priority: Inform about working life	4.53 (2.36)	4.64 (2.20)	-0.11 (0.11)
School priority: Teach children to obey adults	3.65 (2.93)	3.42 (2.49)	0.23* (0.13)
School priority: Cope in a competitive society	3.57 (2.58)	3.62 (2.23)	-0.05 (0.08)
School priority: Select for higher education	1.79 (1.84)	1.80 (1.66)	-0.01 (0.07)

\*  $p < 0.10$ , \*\*  $p < 0.05$ , \*\*\*  $p < 0.01$

Standard errors clustered at the municipal level. Variables are on a 1-9 scale (9 = top priority).

Table 1.B.13: Differences in parents choice protocols  
by treatment status, 1967 cohort

	Graded	Ungraded	Difference
Electives chosen for: child ability	0.62 (0.60)	0.61 (0.53)	0.02 (0.02)
Electives chosen for: child preferences	0.37 (0.61)	0.39 (0.53)	-0.02 (0.02)
Electives chosen for: more choice in HS	0.37 (0.61)	0.38 (0.53)	-0.00 (0.02)
Electives chosen for: entrance requirements	0.24 (0.53)	0.26 (0.48)	-0.01 (0.01)
Electives chosen for: teacher suggestion	0.04 (0.24)	0.07 (0.28)	-0.03*** (0.01)
Electives chosen for: classmates choice	0.01 (0.12)	0.01 (0.13)	-0.00 (0.00)
Electives chosen for: other	0.01 (0.12)	0.01 (0.12)	-0.00 (0.00)

\*  $p < 0.10$ , \*\*  $p < 0.05$ , \*\*\*  $p < 0.01$

Standard errors clustered at the municipal level. Variables represent agreement with the statement on a 0-1 scale.

## 1.C Refutability Tests

### 1.C.1 Tests for Parallel Trends

The following Figures show tests for parallel trends between treatment and control municipalities in determinants of education and educational attainment. For each outcome I plot in the upper panel trends for the grading and non-grading municipalities that are part of my sample. The two dashed lines mark the years in which the 1967 and 1972 cohorts were in school year 6: the tests close to this period are thus more relevant. In the lower panel I show coefficient and 95% confidence interval from difference in differences placebo regressions. In the regressions I control for a linear trend, and run tests over a 5-year window centered on the year marked in the picture. This way I test precisely the assumption that underlies my specification: over a 5-year window there should be no differential trend in education (or related variables) between treated and control municipalities.<sup>45</sup>

In Figure 1.C.1 and Table 1.C.1, I test for parallel trends in the aggregate vote share of right wing parties in municipal elections, held every 4 years in Sweden. It is reassuring to see that the differences in vote share, which can be considered as the “treatment assignment,” are quite stable over time. In Figures 1.C.2 to 1.C.4 I consider educational attainment in the population aged 38-74, corresponding to cohorts who completed their education before the first grading reform, likely including also the parents of the students in the sample. This avoids picking up any effect of the reform. Trends appear to be parallel for all education levels, while there seems to be some catching up on the part of the graded municipalities in high school attainment. That coefficient however is small and marginally significant only in 1989, after the period I consider.<sup>46</sup> Figures 1.C.5 to 1.C.7 consider flows in educational attainment

---

<sup>45</sup>Difference in differences is functional form dependent (Lechner, 2011), and the functional form assumed for the trend should be consistent with the data (Mora & Reggio, 2012). In my analysis I can only control for a linear trend, but this should not be problematic as I use cohorts just 5 years apart.

<sup>46</sup>Notice that administrative education data is only available starting from 1985.



for cohorts born 1969 onwards. These cohorts were all studying under the reformed school system, and thus they started getting grades in school year 8. The first of the pictures plots the fraction of students who graduate from a 2-year high school, which up to the early 90s corresponds to vocational high school.<sup>47</sup> The fraction of students who completes vocational education is higher in grading municipalities, but the diff-in-diff coefficients in the tests are all close to 0, implying that differences remained stable over time. This is one of the key education variables I will be using as an outcome in my analysis, hence it is particularly reassuring to see that the test passes. A similar picture emerges for high school attainment and college.<sup>48</sup>

Table 1.C.1: Test for parallel trends in pro-grade vote share

Vote share of pro-grades parties (%)	
<b>Graded x Year</b>	-0.05 (0.08)
Graded	111.89 (160.44)
Year	-24.62 (25.11)
Mean	35.23
$R^2$	0.32

Standard errors clustered at the municipality level.

*Note:* \*  $p < 0.10$ , \*\*  $p < 0.05$ , \*\*\*  $p < 0.01$

<sup>47</sup>A reform in the 90s increased the length of vocational training to three years.

<sup>48</sup>Statistics Sweden was changing the classification of education in the last part of the panel, a thus there are breaks in the trends. A dashed line marks the first and last year in which estimates are affected by the break, which may lead to spurious rejections in the tests.

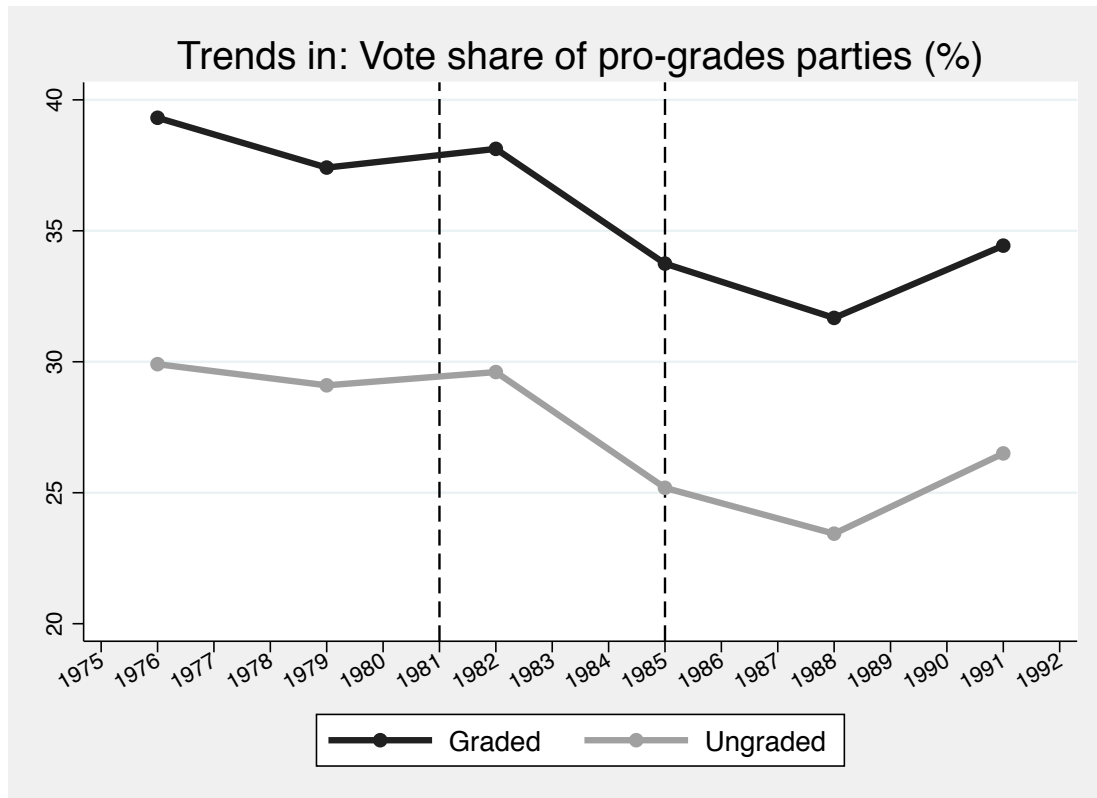


Figure 1.C.1: Test for parallel trends in treatment assignment

Note: The Figure plots aggregate vote share of right-wing parties (in general favoring early grade assignment) in municipal elections, held every 4 years in Sweden, for grading and non-grading municipalities in the sample. Dashed vertical lines mark the period in which the two cohorts in the sample were in school year 6.

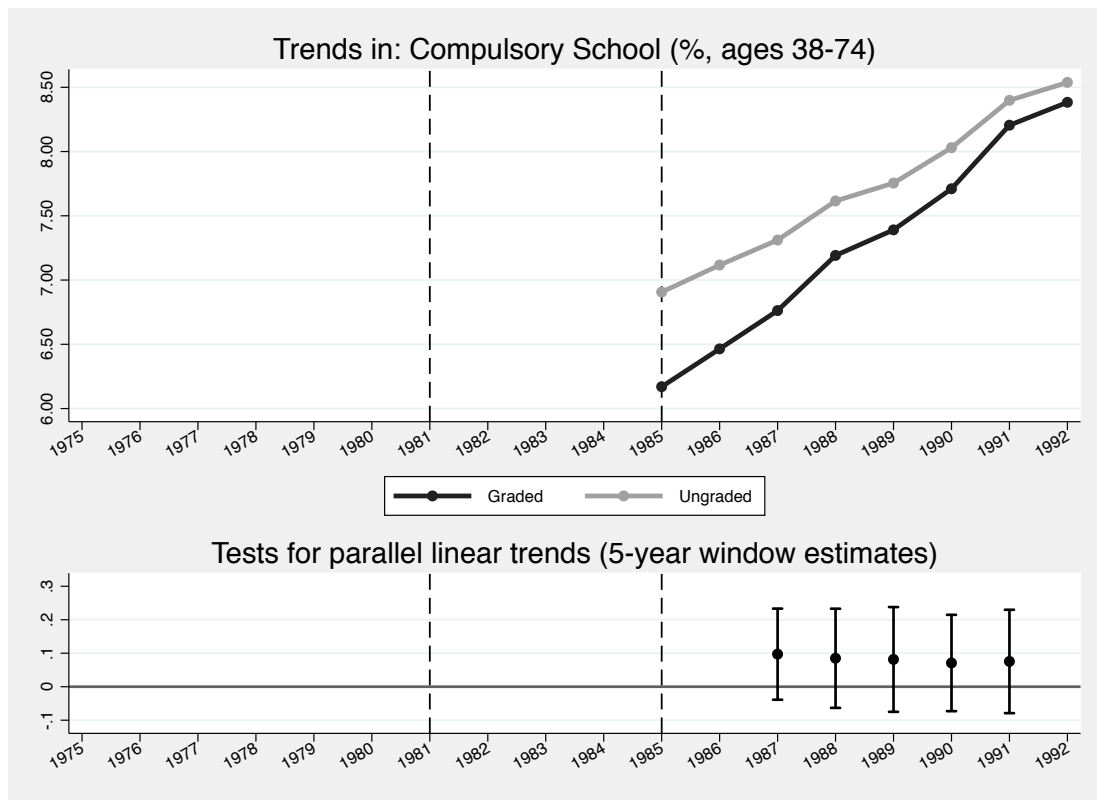


Figure 1.C.2: Test for parallel trends in education:  
Compulsory school (graded cohorts)

Note: The Figure plots in the upper panel trends in 9-year-compulsory-school attainment for people aged 38-74, who studied before the initial reform, for grading and non-grading municipalities in the sample. Dashed vertical lines mark the period in which the two cohorts in the sample were studying. The lower panel presents coefficient and 95% confidence interval from diff-in-diff placebo regressions. Tests are run over a 5-year window centered on the marked year, and control for a linear trend.

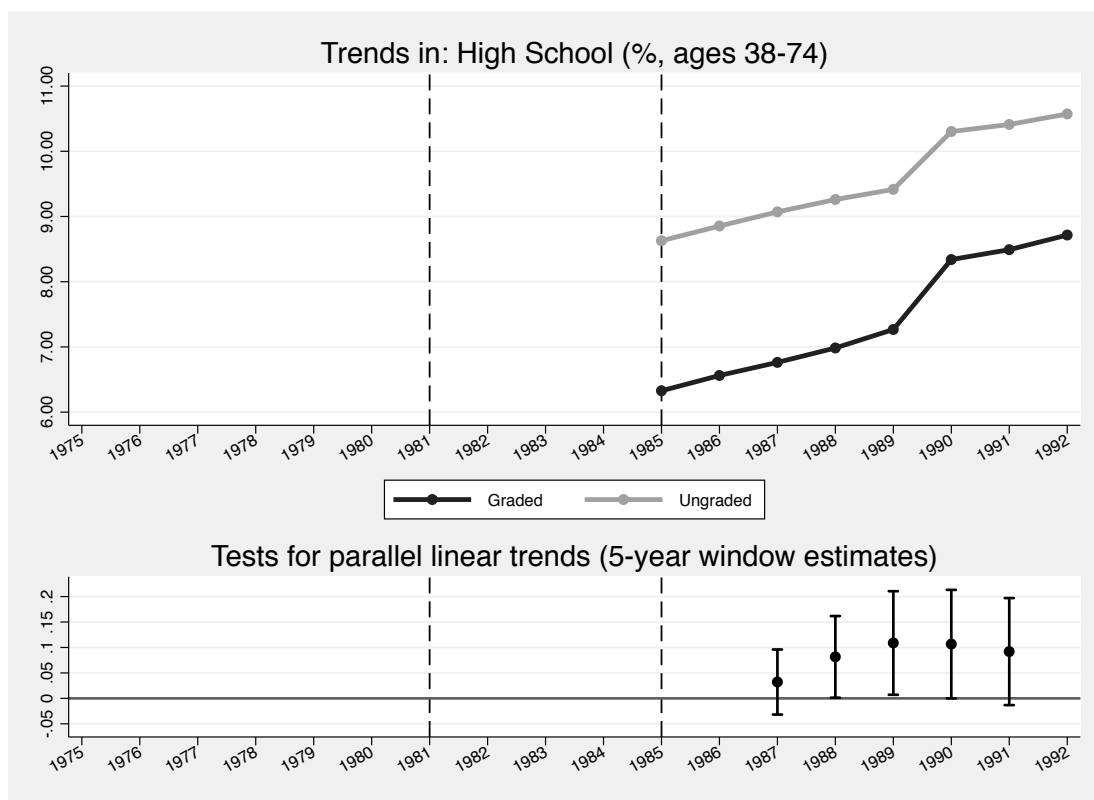


Figure 1.C.3: Test for parallel trends in education:  
High school (graded cohorts)

Note: The Figure plots in the upper panel trends in high-school attainment for people aged 38-74, who studied before the initial reform, for grading and non-grading municipalities in the sample. Dashed vertical lines mark the period in which the two cohorts in the sample were studying. The lower panel presents coefficient and 95% confidence interval from diff-in-diff placebo regressions. Tests are run over a 5-year window centered on the marked year, and control for a linear trend.

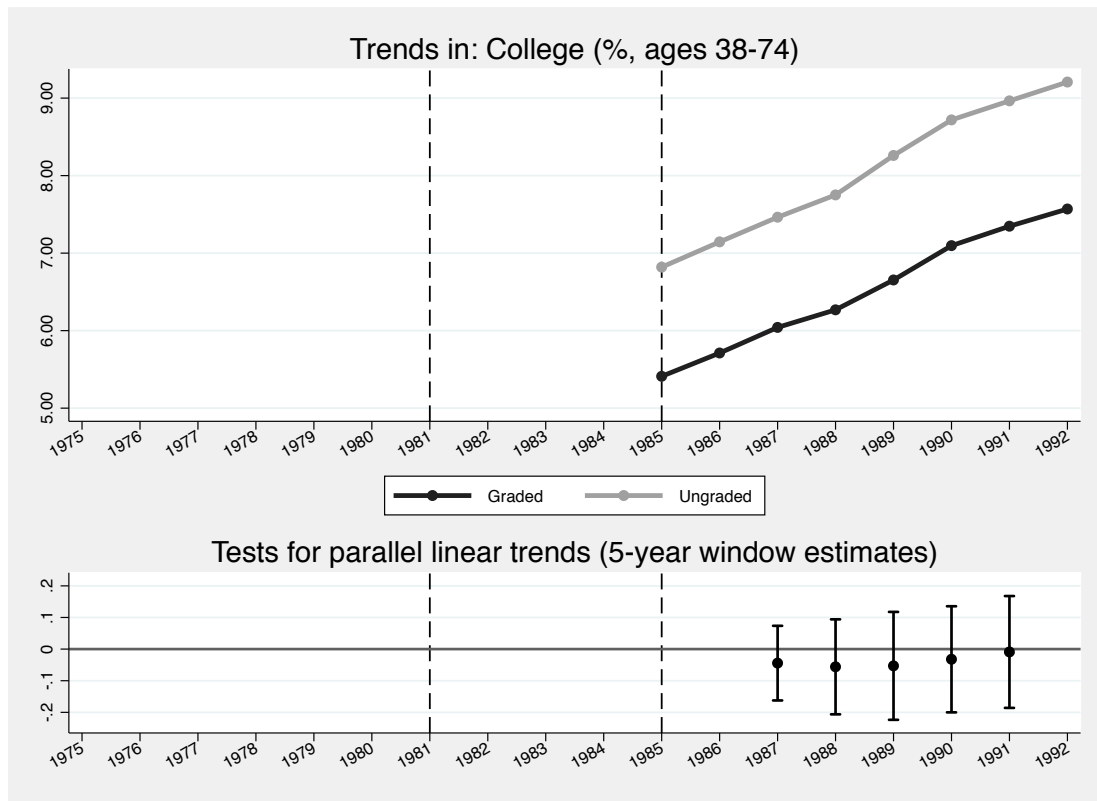


Figure 1.C.4: Test for parallel trends in education:  
College (graded cohorts)

Note: The Figure plots in the upper panel trends in college attainment for people aged 38-74, who studied before the initial reform, for grading and non-grading municipalities in the sample. Dashed vertical lines mark the period in which the two cohorts in the sample were studying. The lower panel presents coefficient and 95% confidence interval from diff-in-diff placebo regressions. Tests are run over a 5-year window centered on the marked year, and control for a linear trend.

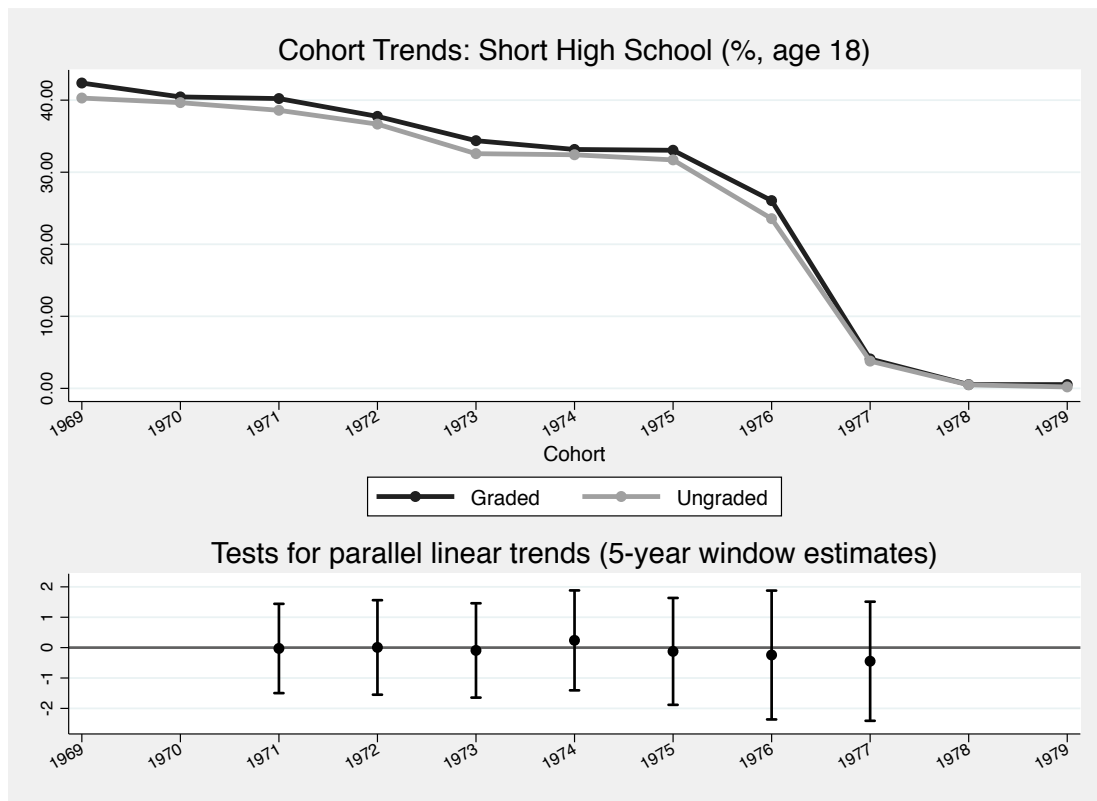


Figure 1.C.5: Test for parallel trends in education flows:  
Short HS (ungraded cohorts)

Note: The Figure plots in the upper panel trends in vocational high school attainment for cohorts who studied when early grades were abolished, for grading and non-grading municipalities in the sample. The lower panel presents coefficient and 95% confidence interval from diff-in-diff placebo regressions. Tests are run over a 5-year window centered on the marked year, and control for a linear trend.

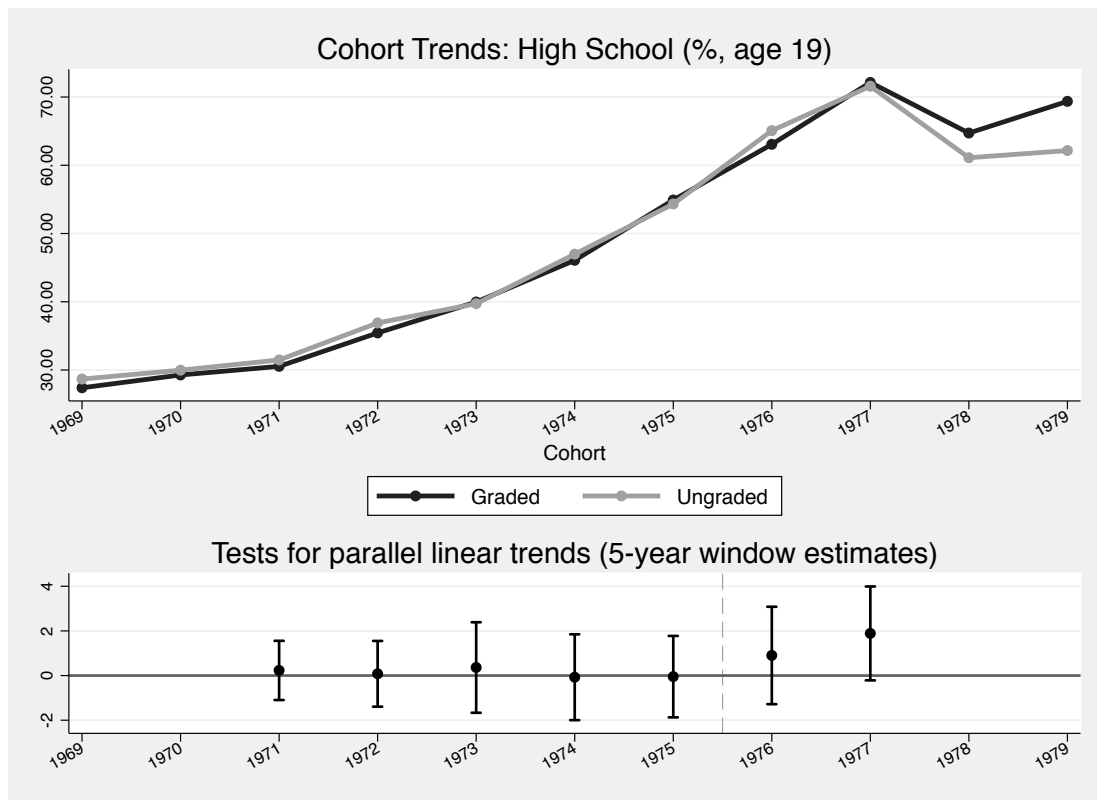


Figure 1.C.6: Test for parallel trends in education flows:  
High school (ungraded cohorts)

Note: The Figure plots in the upper panel trends in high school attainment for cohorts who studied when early grades were abolished, for grading and non-grading municipalities in the sample. The lower panel presents coefficient and 95% confidence interval from diff-in-diff placebo regressions. Tests are run over a 5-year window centered on the marked year, and control for a linear trend. The dashed line marks a break in the time series due to reclassification of education by Statistics Sweden.

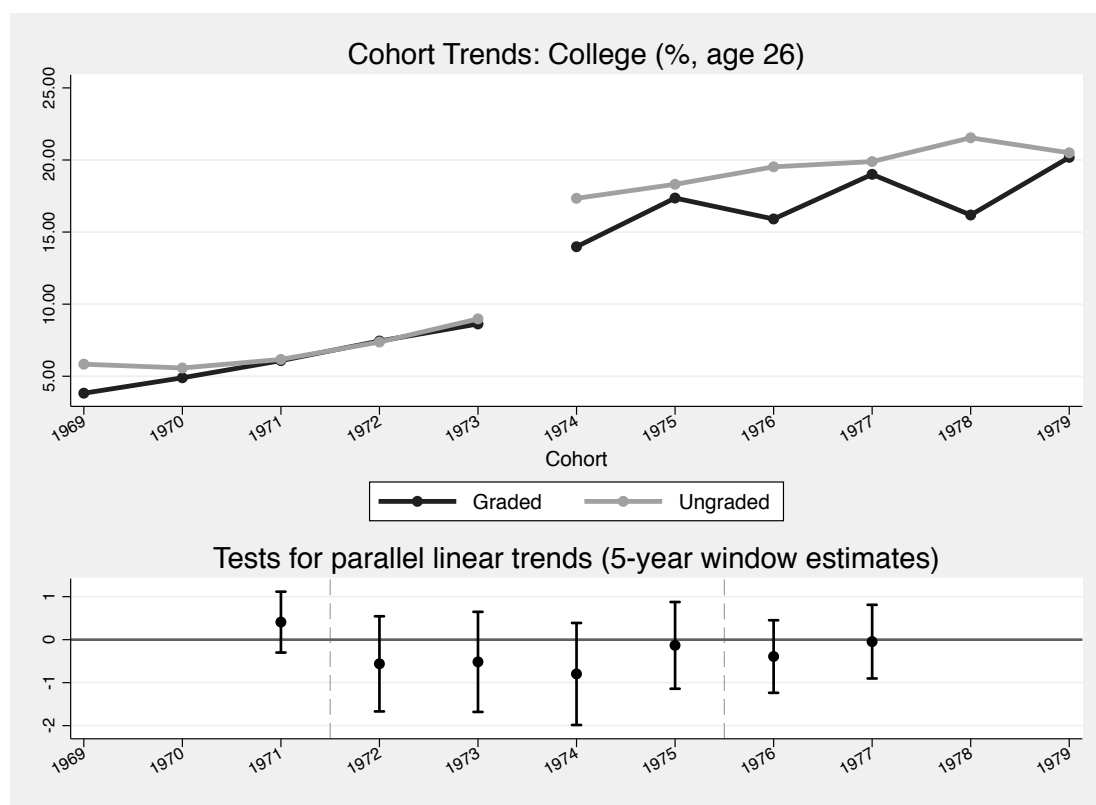


Figure 1.C.7: Test for parallel trends in education flows:  
College (ungraded cohorts)

Note: The Figure plots in the upper panel trends in college attainment for cohorts who studied when early grades were abolished, for grading and non-grading municipalities in the sample. The lower panel presents coefficient and 95% confidence interval from diff-in-diff placebo regressions. Tests are run over a 5-year window centered on the marked year, and control for a linear trend. Estimates within the dashed lines are affected by a break in the time series due to reclassification of education by Statistics Sweden.



### 1.C.2 Tests for Differential Response and Compositional Change

Intelligence and SES data is missing for 18% of my sample, but attrition does not change over time between grading and non-grading municipalities (see Table 1.C.3). It is thus possible to explore heterogeneous effects by ability and SES. There appear to be no issues for the surveys taken in grades 6 and 10, but parental surveys display differential attrition (see Table 1.C.2). Thus I can not use variables from these surveys in the final specification. Finally, among the standardized tests that end-of-the-year grades are based upon, only the Swedish test does not exhibit differential attrition. So this will be the only standardized test I will be using as an outcome.

In Tables 1.C.3 to 1.C.8 I test for differential compositional change in the two sets of municipalities for a large set of pre-treatment variables. As I run the tests for many outcomes and for both the whole sample and the individual ability-SES cells I am likely to find spurious rejections. I thus comment on how the tests perform on average.<sup>49</sup> All demographic and school-level variables pass the tests (see Tables 1.C.4 and 1.C.5). The placebo tests for relative verbal and inductive ability fail in some cases within cell, but are by definition 0 in the sample since I normalized ability at the treatment cohort level.<sup>50</sup> The cross-sectional differences in marriage and divorce rates found for parents seem to persist over time (Tables 1.C.6 and 1.C.7). Both income (Table 1.C.5) and a broad measure of parental education (SES, in Table 1.C.3) pass the tests. Finally most of the occupational categories (Table 1.C.8) pass the tests for compositional change, confirming that the cross sectional differences in occupation remained constant over time. When looking at parental educational attainment, it appears that the fraction of students with college educated mothers (Table 1.C.6) and fathers (Table 1.C.7) increases less in early grading municipalities with respect to the late grading municipalities. These differences are consistent with the two tests that fail for compositional change in occupation: the share of parents involved in non-manual occupations increased less over time in early grading municipalities.

---

<sup>49</sup>In some cases the tests pass in the sample, but not within ability-SES cell. This could be due to compositional change that I find in the ability measures used for the cell. In my analysis I thus always control for ability.

<sup>50</sup>See the discussion in Appendix 1.B.1.

Table 1.C.2: Tests for differential response:  
Summary of difference in differences estimates

Outcome:	Low-ability			High-ability	
	All Sample	Low-SES	High-SES	Low-SES	High-SES
Student survey (school year 6)	-0.03 (0.73) [0.90]	0.01 (0.50) [0.99]	0.01 (0.16) [0.99]	0.01 (0.83) [0.99]	0.01 (0.56) [0.99]
Student survey (school year 10)	0.02 (0.26) [0.76]	0.02 (0.37) [0.69]	0.04* (0.05) [0.77]	0.05** (0.02) [0.83]	0.03 (0.19) [0.88]
Parent survey	-0.06 (0.15) [0.74]	-0.01 (0.67) [0.71]	-0.06** (0.05) [0.78]	-0.01 (0.81) [0.82]	-0.05*** (0.01) [0.87]
English Test (school year 8)	0.13* (0.08) [0.24]	0.10 (0.20) [0.27]	0.09 (0.31) [0.21]	0.17* (0.09) [0.20]	0.14 (0.13) [0.18]
Swedish Test (school year 9)	0.08 (0.45) [0.33]	0.10 (0.34) [0.34]	0.07 (0.56) [0.32]	0.12 (0.34) [0.30]	0.07 (0.52) [0.30]
Math Test (school year 9)	-0.26** (0.02) [0.27]	-0.29*** (0.00) [0.30]	-0.26** (0.02) [0.25]	-0.28*** (0.01) [0.23]	-0.31*** (0.00) [0.22]

\*  $p < 0.10$ , \*\*  $p < 0.05$ , \*\*\*  $p < 0.01$

Wild Cluster Bootstrap p-values in parentheses; sample averages in brackets.

Table 1.C.3:  
Tests for compositional change:  
SES and ability.  
Summary of diff-in-diff estimates

	All Sample
No SES or ability data	0.03 (0.64) [0.18]
High-SES	0.00 (0.84) [0.44]
Low-ability Low-SES	-0.00 (0.87) [0.38]
Low-ability High-SES	-0.04* (0.07) [0.21]
High-ability Low-SES	0.01 (0.55) [0.17]
High-ability High-SES	0.03** (0.04) [0.23]

\*  $p < 0.10$ , \*\*  $p < 0.05$ , \*\*\*  $p < 0.01$

Wild Cluster Bootstrap p-values in parentheses; sample averages in brackets.

Table 1.C.4: Tests for compositional change: Demographics.  
Summary of difference in differences estimates

Outcome:	Low-ability			High-ability	
	All Sample	Low-SES	High-SES	Low-SES	High-SES
Female	-0.02 (0.52) [0.49]	-0.03 (0.50) [0.47]	-0.04 (0.19) [0.48]	0.01 (0.71) [0.52]	0.04 (0.42) [0.50]
Birth year	0.01 (0.53) [1,969.39]	0.01 (0.36) [1,969.20]	-0.01 (0.74) [1,969.65]	0.01 (0.19) [1,969.18]	0.02** (0.03) [1,969.63]
Verbal ability	0.01 (0.85) [-0.00]	0.02 (0.54) [-0.53]	-0.09* (0.07) [-0.33]	-0.10 (0.13) [0.60]	-0.06 (0.32) [0.77]
Inductive ability	0.01 (0.86) [0.00]	0.00 (0.95) [-0.53]	0.02 (0.60) [-0.41]	-0.07 (0.23) [0.70]	-0.12*** (0.00) [0.75]
Spatial ability	0.00 (0.91) [0.00]	0.02 (0.47) [-0.33]	-0.06 (0.13) [-0.15]	-0.08 (0.10) [0.34]	-0.05 (0.25) [0.45]
Special Education	-0.01 (0.68) [0.16]	-0.01 (0.72) [0.25]	-0.04 (0.44) [0.21]	0.02 (0.30) [0.03]	0.02 (0.36) [0.03]

\*  $p < 0.10$ , \*\*  $p < 0.05$ , \*\*\*  $p < 0.01$

Wild Cluster Bootstrap p-values in parentheses; sample averages in brackets.

Table 1.C.5: Tests for compositional change: School and SES.  
Summary of difference in differences estimates

Outcome:	Low-ability			High-ability	
	All Sample	Low-SES	High-SES	Low-SES	High-SES
Switched Class	-0.00 (0.98) [0.12]	0.01 (0.75) [0.09]	-0.02 (0.41) [0.14]	0.01 (0.78) [0.07]	-0.00 (1.00) [0.12]
Changes of teacher	0.03 (0.80) [0.55]	0.04 (0.75) [0.59]	0.14 (0.37) [0.53]	0.03 (0.81) [0.52]	0.09 (0.54) [0.46]
Hours absent (s.y. 6)	-0.79 (0.86) [26.19]	-2.65 (0.55) [26.16]	2.26 (0.59) [26.40]	-0.16 (1.00) [24.15]	2.13 (0.65) [25.98]
Class size	0.44 (0.50) [21.39]	0.16 (0.89) [21.14]	1.08 (0.31) [21.54]	0.24 (0.71) [21.62]	0.60 (0.40) [21.58]
Parents not Nordic	0.00 (0.86) [0.02]	-0.02** (0.03) [0.02]	0.02 (0.11) [0.02]	0.00 (0.90) [0.01]	0.01 (0.44) [0.01]
Quiet home envir.	-0.00 (0.75) [0.95]	-0.02 (0.46) [0.93]	-0.00 (0.87) [0.95]	0.01 (0.35) [0.96]	0.02 (0.27) [0.96]
Family income	0.42 (0.36) [272.37]	0.30 (0.46) [245.29]	0.26 (0.66) [297.31]	0.30 (0.68) [248.11]	-0.02 (0.98) [316.44]

\*  $p < 0.10$ , \*\*  $p < 0.05$ , \*\*\*  $p < 0.01$

Wild Cluster Bootstrap p-values in parentheses; sample averages in brackets. Family income measured when the child is 11-16.

Table 1.C.6: Tests for compositional change: Mother.  
Summary of difference in differences estimates

Outcome:	Low-ability			High-ability	
	All Sample	Low-SES	High-SES	Low-SES	High-SES
Not Nordic	0.01 (0.43) [0.05]	-0.01 (0.21) [0.05]	0.01 (0.49) [0.05]	0.00 (0.87) [0.03]	0.02 (0.11) [0.04]
Married	-0.01 (0.52) [0.74]	-0.01 (0.54) [0.72]	0.03 (0.27) [0.75]	0.00 (0.95) [0.74]	-0.04 (0.20) [0.79]
Divorced	-0.01 (0.44) [0.18]	-0.02 (0.40) [0.17]	-0.04* (0.07) [0.19]	-0.03** (0.04) [0.15]	0.01 (0.78) [0.17]
Compulsory or less	-0.03 (0.14) [0.35]	-0.01 (0.76) [0.53]	-0.07*** (0.00) [0.16]	-0.02 (0.62) [0.48]	-0.07*** (0.00) [0.11]
High school	-0.01 (0.82) [0.44]	-0.02 (0.64) [0.47]	0.04* (0.10) [0.43]	-0.02 (0.61) [0.52]	0.04 (0.42) [0.36]
College or more	0.04* (0.06) [0.21]	0.03** (0.05) [0.00]	0.03 (0.28) [0.41]	0.03* (0.05) [0.00]	0.02 (0.69) [0.53]
In the labor force	-0.00 (0.89) [0.89]	0.00 (0.91) [0.86]	0.02 (0.19) [0.93]	0.01 (0.65) [0.88]	-0.03 (0.34) [0.95]
Unemployed	0.00 (0.47) [0.05]	0.01 (0.26) [0.06]	0.01 (0.37) [0.04]	-0.00 (0.99) [0.05]	0.00 (0.81) [0.04]

\*  $p < 0.10$ , \*\*  $p < 0.05$ , \*\*\*  $p < 0.01$

Wild Cluster Bootstrap p-values in parentheses; sample averages in brackets.

Table 1.C.7: Tests for Compositional Change: Father.  
Summary of difference in difference estimates

Outcome:	Low-ability			High-ability	
	All Sample	Low-SES	High-SES	Low-SES	High-SES
Not Nordic	0.01 (0.53) [0.03]	-0.01 (0.32) [0.04]	0.02* (0.10) [0.03]	0.00 (0.69) [0.02]	0.01 (0.37) [0.03]
Married	0.00 (0.78) [0.77]	-0.00 (0.98) [0.75]	0.00 (0.88) [0.78]	-0.01 (0.53) [0.79]	-0.02 (0.30) [0.82]
Divorced	-0.01 (0.57) [0.17]	0.01 (0.74) [0.17]	-0.02 (0.16) [0.17]	-0.00 (0.91) [0.14]	-0.00 (0.85) [0.15]
Compulsory or less	-0.01 (0.32) [0.41]	-0.01 (0.64) [0.67]	-0.01 (0.78) [0.13]	0.03 (0.39) [0.63]	-0.05** (0.03) [0.09]
High school	-0.02 (0.20) [0.40]	-0.02 (0.50) [0.33]	0.01 (0.67) [0.51]	-0.06* (0.09) [0.37]	0.02 (0.57) [0.44]
College or more	0.04*** (0.00) [0.19]	0.03** (0.04) [0.00]	-0.00 (0.87) [0.36]	0.03** (0.04) [0.00]	0.03 (0.22) [0.47]
In the labor force	-0.01 (0.24) [0.88]	0.01 (0.70) [0.85]	-0.02 (0.33) [0.91]	-0.03 (0.32) [0.88]	-0.02* (0.06) [0.93]
Unemployed	0.00 (0.60) [0.04]	-0.00 (0.91) [0.05]	-0.01 (0.32) [0.04]	0.02 (0.19) [0.04]	0.04** (0.01) [0.04]

\*  $p < 0.10$ , \*\*  $p < 0.05$ , \*\*\*  $p < 0.01$

Wild Cluster Bootstrap p-values in parentheses; sample averages in brackets.

Table 1.C.8: Tests for Compositional Change:  
Best of parent occupation.  
Summary of difference in difference estimates

	Diff-in-Diff (p-value)
Parents: non-skilled workers, goods	-0.01 (0.69)
Parents: non-skilled workers, service	0.01 (0.46)
Parents: skilled workers, goods	-0.06*** (0.00)
Parents: skilled workers, service	0.01 (0.17)
Parents: lower non-manual ii	0.00 (0.96)
Parents: lower non-manual i	0.03*** (0.00)
Parents: intermediate-level non-manual	0.02 (0.45)
Parents: higher civil servants and senior salaried	0.02 (0.21)
Parents: independent professionals	-0.00* (0.09)
Parents: entrepreneur	-0.01 (0.42)
Parents: farmer	-0.00 (0.60)

\*  $p < 0.10$ , \*\*  $p < 0.05$ , \*\*\*  $p < 0.01$

Standard errors clustered at the municipal level. Occupation variables are taken from Census 1980.



## References

- Altonji G. Joseph, 1993. "The Demand for and Return to Education When Education Outcomes Are Uncertain," *Journal of Labor Economics*, University of Chicago Press, vol. 11(1), pages 48-83, January.
- Altonji G. Joseph & Blom Erica & Meghir Costas, 2012. "Heterogeneity in Human Capital Investments: High School Curriculum, College Major, and Careers," *NBER Working Papers* 17985, National Bureau of Economic Research, Inc.
- Angrist D. Joshua & Pischke Jörn-Steffen. "Mostly Harmless Econometrics: An Empiricist's Companion," Princeton university press, 2008.
- Arcidiacono Peter, 2004. "Ability Sorting and the Returns to College Major," *Journal of Econometrics*, Elsevier, vol. 121(1-2), pages 343-375.
- Arellano Manuel, 1987 "Computing Robust Standard Errors for Within-Group Estimators," *Oxford Bulletin of Economics and Statistics* 49, 431-434.
- Azmat Ghazala & Iriberry Nagore, 2010. "The Importance of Relative Performance Feedback Information: Evidence from a Natural Experiment using High School Students," *Working Papers* 444, Barcelona Graduate School of Economics.
- Bandiera Oriana & Larcinese Valentino & Rasul Imran, 2008. "Blissful Ignorance? Evidence from a Natural Experiment on the Effect of Individual Feedback on Performance," mimeo.
- Becker S. Gary , 1994. "Human Capital: A Theoretical and Empirical Analysis with Special Reference to Education," 3rd Edition, *NBER Books*, National Bureau of Economic Research, Inc, number beck94-1, August.
- Ben-Porath Yoram, 1967. "The Production of Human Capital and the Life Cycle of Earnings," *Journal of Political Economy*, University of Chicago Press, vol. 75, pages 352.
- Betts R. Julian, 1998. "The Impact of Educational Standards on the Level and Distribution of Earnings," *American Economic Review*, American Economic Association, vol. 88(1), pages 266-75, March.

Betts R. Julian & Grogger Jeff, 2003. "The Impact of Grading Standards on Student Achievement, Educational Attainment, and Entry-level Earnings," *Economics of Education Review*, Elsevier, vol. 22(4), pages 343-352, August.

Becker W. & Rosen S., 1990. "The Learning Effect Of Assessment And Evaluation In High School," University of Chicago - Economics Research Center 90-7, Chicago - Economics Research Center.

Bertrand Marianne, 2011. "New Perspectives on Gender," *Handbook of Labor Economics*, Elsevier.

Bertrand Marianne & Duflo Esther & Mullainathan Sendhil, 2004. "How Much Should We Trust Differences-in-Differences Estimates?," *The Quarterly Journal of Economics*, MIT Press, vol. 119(1), pages 249-275, February.

Blundell Richard & Costa Dias Monica, 2009. "Alternative Approaches to Evaluation in Empirical Microeconomics," *Journal of Human Resources*, University of Wisconsin Press, vol. 44(3).

Börklund Anders, 1993. "A Comparison between Actual Distributions of Annual and Lifetime Income: Sweden 1951-89," *Review of Income and Wealth*, International Association for Research in Income and Wealth, vol. 39(4), pages 377-86, December.

Björklund Anders & Clark A. Melissa & Edin Per-Anders & Fredriksson Peter & Krueger B. Alan , 2005. "The Market Comes to Education in Sweden: An Evaluation of Sweden's Surprising School Reforms," Russell Sage Foundation Publications.

Borghans Lex & Lee Duckworth Angela & Heckman James & Ter Weel Bas, 2008. "The Economics and Psychology of Personality Traits," *Journal of Human Resources*, University of Wisconsin Press, vol. 43(4).

Borghans Lex & Meijers Huub & Weel Bas Ter, 2008. "The Role Of Noncognitive Skills In Explaining Cognitive Test Scores," *Economic Inquiry*, Western Economic Association International, vol. 46(1), pages 2-12, 01.

Brunello Giorgio & Checchi Daniele, 2007. "Does School Tracking Affect Equality of Opportunity? New International Evidence," *Economic Policy*, CEPR;CES;MSH, vol. 22, pages 781-861, October.

- Cameron Colin & Gelbach B. Jonah & Miller L. Douglas, 2008. "Bootstrap-Based Improvements for Inference with Clustered Errors," *The Review of Economics and Statistics*, MIT Press, vol. 90(3), pages 414-427, August.
- Cunha Flavio & Heckman James, 2007. "The Technology of Skill Formation," *American Economic Review*, American Economic Association, vol. 97(2), pages 31-47, May.
- Cunha Flavio & Heckman James, 2009. "The Economics and Psychology of Inequality and Human Development," *Journal of the European Economic Association*, MIT Press, vol. 7(2-3), pages 320-364, 04-05.
- Emanuelson Ingemar, 1979. "Utvärdering Genom Uppföljning av Elever - Ett Nytt Individual-statistikprojekt," Stockholm School of Teacher Education, Report 11-1979.
- Figlio N. David & Lucas E. Maurice, 2004. "Do High Grading Standards Affect Student Performance?," *Journal of Public Economics*, Elsevier, vol. 88(9-10), pages 1815-1834, August.
- Härnqvist Kjell, 1994a. "Den Sociala Selektionen till Gymnasieskola och Högskola," Rapport 1994:10, Göteborgs universitet, Institutionen för pedagogik.
- Härnqvist Kjell, 1998. "A Longitudinal Program for Studying Education and Career Development," Report 1998: 01, Göteborgs universitet, Institutionen för pedagogik (1998).
- Hoxby Caroline, 2011. "Program Report," NBER Economics of Education Program, Issue 1, NBER Reporter.
- Jalava Nina & Joensen Schrøter Juanna & Pellas Elin, 2015. "Grades and Rank: Impacts of non-financial Incentives on Test Performance," *Journal of Economic Behavior & Organization*, Elsevier, vol. 115(C), pages 161-196.
- Hussey J. Andrew & Swinton H. Omari, 2011. "Estimating the Ex Ante Expected Returns to College," *American Economic Review*, American Economic Association, vol. 101(3), pages 598-602, May.

Klapp Alli & Cliffordson Christina & Gustafsson Jan-Eric, 2014. "The Effect of Being Graded on Later Achievement: Evidence from 13-year Olds in Swedish Compulsory School," *Educational Psychology* ahead-of-print: 1-19.

Klapp Alli, 2015. "Does Grading affect Educational Attainment? A Longitudinal Study," *Assessment in Education: Principles, Policy & Practice* ahead-of-print: 1-22.

Kolenikov Stanislav & Angeles Gustavo, 2005. "On Reuse of Clusters in Repeated Studies," presented at the American Statistical Association, Minneapolis, MN.

Lechner Michael, 2011. "The Estimation of Causal Effects by Difference-in-Difference Methods," *Foundations and Trends in Econometrics*, now publishers, vol. 4(3), pages 165-224, November.

Lindqvist Erik & Vestman Roine, 2011. "The Labor Market Returns to Cognitive and Noncognitive Ability: Evidence from the Swedish Enlistment," *American Economic Journal: Applied Economics* 3.1: 101-28.

MacKinnon James & G. & Matthew D. Webb, 2014. "Wild Bootstrap Inference for Wildly Different Cluster Sizes," *Working Papers* 1314, Queen's University, Department of Economics.

Meghir Costas & Palme Mårten, 2005. "Educational Reform, Ability, and Family Background," *American Economic Review*, American Economic Association, vol. 95(1), pages 414-424, March.

Mora Ricardo & Reggio Iliana, 2012. "Treatment effect identification using alternative parallel assumptions," *Economics Working Papers* we1233, Universidad Carlos III, Departamento de Economía.

Öckert Björn, 2002. "Do University Enrollment Constraints Affect Education and Earnings?," *Working Paper Series* 2002:16, IFAU - Institute for Evaluation of Labour Market and Education Policy.

Öckert Björn, 2010. "What's the Value of an Acceptance Letter? Using Admissions Data to Estimate the Return to College," *Economics of Education Review*, Elsevier, vol. 29(4), pages 504-516, August.

OECD (2010), "How Many Students Drop Out of Tertiary Education?," in Highlights from Education at a Glance 2010, OECD Publishing.

Régner Isabelle & Pascal Huguet & Jean-Marc Monteil, 2002. "Effects of Socioeconomic Status (SES) Information on Cognitive Ability Inferences: When Low-SES Students Make Use of a Self-threatening Stereotype," *Social Psychology of Education* 5.3: 253-269.

Sjögren Anna, 2010. "Graded children – Evidence of Longrun Consequences of School Grades from a Nationwide Reform," *Working Paper Series* 2010:7, IFAU - Institute for Evaluation of Labour Market and Education Policy.

Spence Michael, 1973. "Job Market Signaling," *The Quarterly Journal of Economics*, 355-374.

Stinebrickner Todd & Stinebrickner Ralph, 2012. "Learning about Academic Ability and the College Dropout Decision," *Journal of Labor Economics*, University of Chicago Press, vol. 30(4), pages 707 - 748.

Stange Kevin, 2012. "An Empirical Investigation of the Option Value of College Enrollment," *American Economic Journal: Applied Economics*, American Economic Association, vol. 4(1), pages 49-84, January.

Svensson Allan, 1995. "Att Välja eller Välja Bort Naturvetenskap och Teknik," NOT-projektet, 3.

Todd Petra & Wolpin Kenneth, 2003. "On the Specification and Estimation of the Production Function for Cognitive Achievement," *The Economic Journal*, 113 (485), F3-F33.

White H., 1984. "Asymptotic Theory for Econometricians," San Diego: Academic Press.

Zafar Basit, 2011. "How Do College Students Form Expectations?," *Journal of Labor Economics*, University of Chicago Press, vol. 29(2), pages 301 - 348.



## Chapter 2

# Rethinking Education Choices: The Effect of Surveys<sup>1</sup>

---

<sup>1</sup>This paper is joint work with Juanna Schrøter Joensen (University of Chicago, Stockholm School of Economics, and IZA) and Gregory Francisco Veramendi (Arizona State University). We thank seminar participants at the Stockholm School of Economics for useful feedback. We gratefully acknowledge financial support from the Swedish Foundation for Humanities and Social Sciences (Riksbankens Jubileumsfond) grant P12-0968 enabling the data collection for this project. The usual disclaimers apply.

## 2.1 Introduction

Empirical work in the social and health sciences depends crucially on the use of survey data. Surveying individuals may draw their attention to risks, returns, or choices previously not salient to them. This may change their subsequent behavior. If so, it may bias parameter estimates and any conclusions drawn from survey data. In this paper, we assess whether surveys can causally affect educational investments. We also assess for *whom* and *how* surveys can change education choices and outcomes.

Does responding to a survey designed to evaluate own performance in school and reflect on future education opportunities cause individuals to re-think their educational choices? We answer this question by linking survey data to comprehensive administrative data for the whole Swedish population. The extensive education survey was administered to a randomly drawn sample of three cohorts of 3rd graders in the 80s and early 90s. The survey thus constitutes a randomized social experiment for testing whether reflecting on survey questions alters behavior. We observe education and labor market outcomes when individuals are 28-31. Importantly, these are from administrative registers and *not* reported by the individuals themselves. This means that measurement error is minimal in our data, we can conduct balancing tests on a rich set of pre-determined characteristics of surveyed and non-surveyed individuals to corroborate the success of random assignment, *and* we can analyze the impact of the survey on both shorter- and longer-term outcomes.

If being surveyed changes behavior, then it has implications for both the external and internal validity of studies based on survey data. The total “treatment” effect of being surveyed is thus interesting per se. Most countries administer Household and Labor Force Surveys (LFS). Much of the research on education, labor market, household finance, health, and retirement choices is based on survey data. For example, the Panel Study of Income Dynamics (PSID) which is a longitudinal household survey which began in 1968 with a nationally representative sample of US households, and the National Longitudinal Survey of Youth 1979 (NLSY79) and 1997 (NLSY97). These surveys comprise samples of the cohorts born in 1957-64 (1980-84) and respondents were aged 14-22 (12-17) when first interviewed in 1979 (1997). The construction of these panel surveys has been crucial for many of the recent methodological advances in microeconomics and applied microeconomics more generally. The empirical issues of



dealing with measurement error and validation have been extensively studied.<sup>2</sup> The fact that repeated surveying can alter individual survey response patterns has also been well-established. For example, related to the monthly Current Population Survey (CPS) rotation group bias and its impact on unemployment estimates (Bailar, 1975; Solon, 1986). However, whether responding to surveys can actually alter behavior is still an open question. The only study we know of (Duflo et al., 2011) finds mixed evidence. Duflo et al. (2011) randomly assign individuals in five field experiments (three on health and two on micro-lending) to respond to survey questions on health and/or household finance. They find that responding to health-related questions significantly altered health-related behavior. Those randomized to take the health survey had significantly higher take-up of medical insurance and increased use of water treatment products. They also find that this leads to biased estimates of their estimated impact of improved water source quality - despite random assignment to higher water source quality. These results indicate that researchers should be cautious when administering extensive and repeated surveys, since they may alter the estimated treatment effects of those surveyed by changing their behavior. However, these results seem to be context-dependent.

This paper differs from Duflo et al. (2011) in five important aspects. First, we merge the random survey sample to administrative data for the whole Swedish population. Our sample size is therefore much larger and the measurement errors in education choices and outcomes are minimal. Second, we analyze the effect of being surveyed in a developed (rather than developing) country. Third, we examine the domain of education (rather than health and lending). This is important if impacts are context-specific. For example, if survey effects only arise in settings where individuals previously ignored some of their potential choices and opportunities (e.g. enrolling in high school or college). In such settings, the survey can make these choices more salient, divert the focus to rethink priors, and spur individuals to take otherwise missed opportunities (e.g. more rewarding educational paths). Fourth, we have access to complete medical birth records and a range of measures of family composition and resources to conduct balancing tests of pre-determined characteristics of surveyed and non-surveyed individuals. Lastly, we are able to follow individuals for 18-21 years after random survey assignment. This allows us to analyze both

---

<sup>2</sup>See e.g. Bound et al. (2001) for an extensive survey of the literature.

shorter- and longer-term outcomes, which is important if the strength of the effect of being surveyed diminishes or amplifies over time.

Surveys may affect education choices through providing information to rethink education choices and potentially change expectations. Our paper is therefore also related to the literature analyzing the effects of information on the returns to investment in education. Jensen (2010) finds that his sample of 8th grade boys in the Dominican Republic significantly underestimate the returns to schooling. Informing a random subset of them about higher measured returns leads to a significant increase in perceived returns and 0.20-0.35 more years of schooling. Similarly, Nguyen (2008) finds that informing a random subset of primary school students in rural Madagascar about the average returns to schooling increased their attendance rates by 3.5 percentage points and their test scores by 0.2 standard deviations on average. Students whose priors were below the informed average return, had an even higher increase in test scores of 0.37 standard deviations. These studies analyze very short term effects of providing specific information on the population distribution of the returns to education on educational attainment four months (Nguyen, 2008) to four years later (Jensen, 2010). Relatedly, Wiswall and Zafar (2015) estimate the effects of providing information on college major-specific characteristics on major choice during the college years, and Jalava et al. (2015) estimate the effects of information on test assessment on immediate test effort and performance. In this paper, we are also able to access longer term effects on completed schooling and realized labor market returns.

The “treatment” of being surveyed is a bundle of different types of information. This means that there are several potential channels through which the survey can affect education choices. We try to disentangle these channels exploiting variation in respondents, questions, and timing of questions. Three potential channels are: first, the children need to evaluate themselves and their abilities - also relative to their peers’ abilities. They also take one or multiple cognitive aptitude tests. Even if they are not informed of their test scores, the test situation may still convey information to those in the tails; e.g. if they could not reach the end of the test, were not able to answer many test items, or finished before their peers and confidently solved all test items. Second, the survey required them to state their preferences; e.g. their desired occupation and spell out their future education plans. This can be seen as a “nudge” to rethink and evaluate goals (i.e. desired education and jobs), means, and costs in

a forward-looking manner.<sup>3</sup> This could lead to more well-considered choices and less “mismatch” between individual abilities, education, and career choices.<sup>4</sup> Third, the survey could increase awareness or the salience of choices not previously considered. The education system and institutional setting may not be clear to the child and the parents. Particularly, the connection between early academic choices and the tracked school system. There may also be an informational asymmetry in that parents with higher education may be better informed than parents who dropped out of school after compulsory schooling. The survey asks for reflection on these education choices, including continuing to high school and college. This information could affect choices by increasing awareness of choices not previously considered; especially for parents who never themselves took these educational paths. If so, this has substantive implications for how to model educational choices. Limited attention models, where individuals simplify complex choice settings by only considering a limited subset of choices, have existed at least since Simon (1955).<sup>5</sup> However, such models have not yet been considered in the context of human capital accumulation and education choice. Fourth, parents also respond to questions about school inputs, school choices, and how much of their time they devote to their children – in particular to their schooling investments. This could lead parents to invest more time in their children’s skill accumulation and schooling.<sup>6</sup>

---

<sup>3</sup>In the survey wave the year after compulsory schooling completion (10th grade) they are also asked how they made their education choices of elective courses and how they decided to enroll in high school or not. This is also “nudging” them to evaluate the optimality of their past choices. Some who may otherwise have dropped out after compulsory schooling may thus decide to enroll in high school after having stopped out of school for a year or two.

<sup>4</sup>Several studies find considerable uncertainty about abilities and room for learning through grades and other feedback in education (Altonji, 1993; Arcidiacono, 2004; Arcidiacono et al., 2011, 2012; Zafar, 2011; Facchinello, 2016). More information on educational tracks may also improve outcomes in terms of more sorting on test scores across high school tracks and less dropout (Goux et al., 2014).

<sup>5</sup>See e.g. Barberis and Thaler (2003) and DellaVigna (2009) for reviews on contexts in finance and economics where limited attention has been found important.

<sup>6</sup>More parental involvement in their child’s schooling is found to improve their child’s, and even their child’s peers’, school attendance (Avvisati et al., 2014). Cunha et al. (2010) find that measured parental investments account for 15% of the variation in educational attainment. Heckman and Mosso (2014) provide a recent comprehensive review of this emerging

We shed light on the potential channels by assessing heterogeneity in treatment effects estimated under different identifying assumptions (a within-municipality, a within-school, and two between-school estimators) across subgroups with different levels of parental education.

The causal effect of being surveyed on both short- and long-run outcomes is generally not significantly different from zero, independently of parental education. We find, however, that being surveyed increases educational attainment and job stability in the early career for some subpopulations. We thus think it is worth investigating in future research whether being surveyed affects education choices for specific subpopulations. We will try to disentangle the mechanisms by examining whether it makes a difference when and who is surveyed - in 6th grade or also in 3rd grade, the children themselves or also their parents. First, we have variation in when children responded to the survey, whether and when parents also responded to the survey, and in the type and intensity of survey questions. Second, we are able to merge data on siblings in order to estimate potential information spillover effects. Siblings who were not directly affected by the survey, could only be affected if there are information spillovers through social interaction with their siblings and parents – a potentially important determinant of educational choice (Joensen and Nielsen, 2015). We will exploit that the parent is surveyed when their non-surveyed children (i.e. the siblings of the surveyed child) are at different stages of their educational paths. Some older siblings may be about to make critical decisions on whether to enroll in high school or in college. Therefore, the importance of education may become particularly salient for these siblings as the parent reflects on career choices.

The rest of the paper proceeds as follows: the next section spells out the details of the institutional setting, survey sampling scheme, and how we exploit these in our empirical strategy. Section 2.3 presents the data, descriptive statistics, and balancing tests. Section 2.4 presents the results, while Section 2.5 concludes.

---

literature.

## 2.2 Institutional Setting and Empirical Strategy

This section first provides some background on the institutional setting and the Swedish schooling system the surveyed individuals and their cohorts were facing. Second, we describe the sampling scheme of the Evaluation Through Follow-up (ETF) cohort-sequential longitudinal survey. Third, we describe the empirical strategy we use to identify the effect of surveys on education choice.

### 2.2.1 Survey Sampling Scheme

The ETF survey was administered by the Department of Education and Special Education, Gothenburg University.<sup>7</sup> The survey was constructed through a multistage sampling scheme with stratification at the municipal level: (1) systematically draw two municipalities at random per stratum (13 strata) plus the three largest municipalities (Stockholm, Gothenburg, and Malmö), (2) randomly draw classes within sampled municipalities, and (3) survey all students in each sampled class.

More specifically, the survey sample selection design was as follows. In stage (1) the three largest municipalities (Stockholm, Gothenburg, and Malmö) were selected with probability one, while the remaining 281<sup>8</sup> Swedish municipalities were categorized into 13 strata according to their population (above or below 25,000 inhabitants), proportion of “socialist” mandates (above or below 50%), the share employed in public administration (above or below 25%), and the proportion of immigrant pupils (above or below 8%).<sup>9</sup> Within each stratum, the municipalities were assigned a sampling probability weight,  $p_m$ , proportional to their share of pupils in the relevant school cohort. Finally, two municipalities were sampled at random (conditional on  $p_m$ ) from each stratum. Thus, a total of 29 municipalities were sampled and larger municipalities were more likely to be sampled. Figure 2.2.1 displays a map of Sweden with each of the sampled municipalities in stage (1) for each of the three cohorts.

In stage (2) Statistic Sweden’s (SCB) class register was used to randomly sample classes: 3rd graders in the school-years 1981/82 (ETF72), 1986/87 (ETF77),

<sup>7</sup>Härnqvist (1998) provides additional details on the construction of the survey.

<sup>8</sup>283 in the last wave.

<sup>9</sup>The exact procedure of selecting municipalities is extensively documented in Emanuelsson (1979).

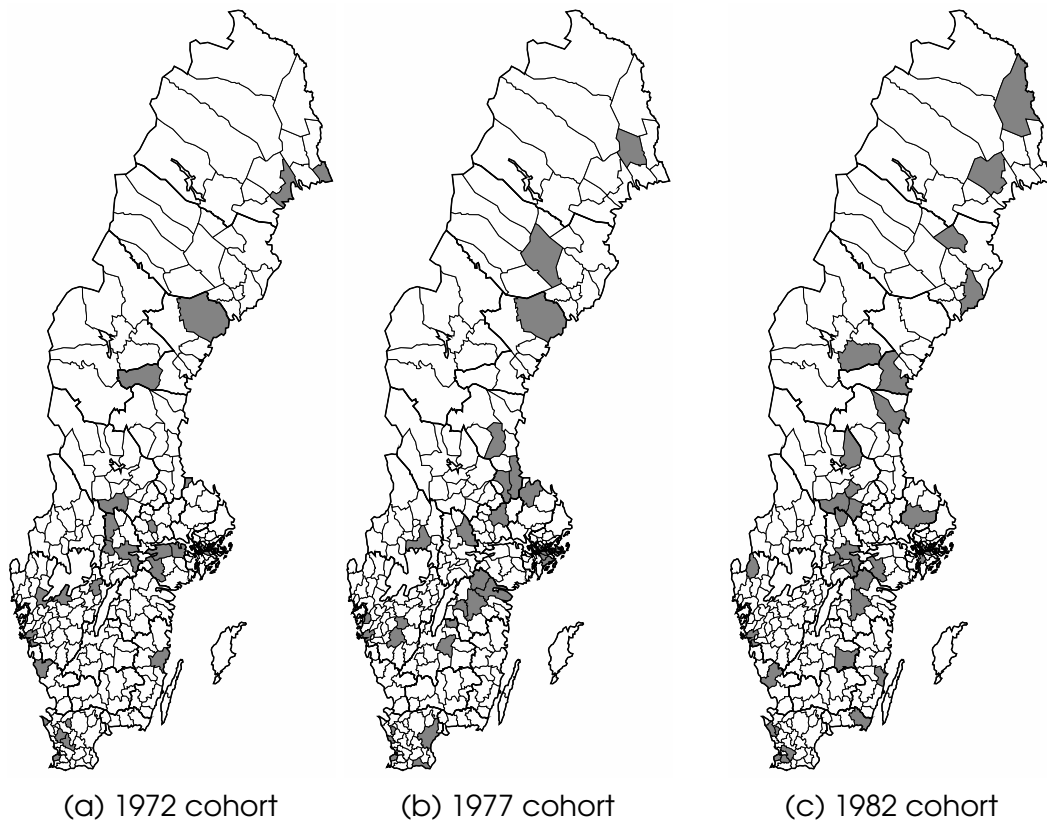


Figure 2.2.1: Sampled municipalities

Note: The Figure displays a map of Sweden with the sampled municipalities in each ETF-cohort 1972, 77, and 82, respectively, shaded.

and 1991/92 (ETF82). Unfortunately, these class registers have not been kept in SCB's archives. This means that we can not perfectly measure "treatment" and "control" group assignment, since we can not perfectly measure who is in the "control" group: i.e. those who attended 3rd grade in the sampled municipalities in the same year, but were not in a sampled class. The sample selection criteria for the classes were that: (2.i) 17 classes in each sampled municipality were selected at random, but with the two exceptions that (2.ii) all classes in small municipalities with 17 or fewer classes were sampled and (2.iii) 29 classes were selected at random in the three largest municipalities (Stockholm, Gothenburg, and Malmö). The class sampling was made in January in the relevant school-year (1981/82, 1986/87, and 1991/92). In stage (3), all students registered in the sampled classes were surveyed on April 15 in the relevant school-year, hence "treated". The first two cohorts (ETF72 and ETF77) were surveyed in 3rd, 6th,

and 10th grade (the year after compulsory schooling was finalized) whereas the last cohort (ETF82) was only surveyed in 6th grade.<sup>10</sup> The students' parents were also surveyed once: in 3rd grade for the first cohort (ETF72) and in 6th grade for the last two cohorts (ETF77 and ETF82). Table 2.2.1 provides an overview of the administered surveys and tests for each cohort.

Table 2.2.1: Treatment Assignment Overview

	Parent survey		Child survey			Aptitude test	
	3rd	6th	3rd	6th	10th	3rd	6th
1972 cohort	T		T	T	T	T	T
1977 cohort		T	T	T	T		T
1982 cohort		T		T		T	

The Table displays an overview of the variation in treatment assignment over grades for each of the three ETF-cohorts 1972, 77, and 82.

## 2.2.2 Control and Treatment Group Assignment

With the data we have, we can make two approximations of the “control” group.

First, those who were in 9th grade in 1987/88, 1992/93, 1997/98. The two main drawbacks with this selection method is that it assumes no grade retention and no students selecting in and out of classes during the six-year period from grade 3 to grade 9. This seems to be an especially problematic assumption for those not born in Sweden, since the surveyed (i.e. the “treatment” group) will not contain any students who have immigrated after grade 3, but the “control group” will. Also students who may re-take or skip a grade will be misclassified according to this assumption. However, grade retention was extremely rare for these cohorts. For the last two cohorts born in 1977 and 1982, we also have information on the municipality of birth from the medical birth registry as well as the municipality of residency of the parents in December 1990 (when the ETF77 cohort attended 7th grade) and in December 1991 (when the ETF82 cohort attended 3rd grade and were sampled). This allows us to

<sup>10</sup>Note that randomization also happened in 3rd grade for the ETF82 cohort despite them not being surveyed until 6th grade.

almost perfectly measure who was in the “control” group in the ETF82 cohort and more accurately measure who was in the “control” group in the ETF77 cohort.

Second, those who are born in 1972, 1977, and 1982. This drops everyone from the “control group” who started school a year “early” or “late” compared to the ordained schooling track. That is all the students not born in these three years, but who still were part of the three relevant schooling cohorts.

We discuss additional sample selection criteria and refinements to the approximation of treatment and control groups in Sections 2.2.3 and 2.3.

### 2.2.3 Empirical Strategy

We now turn to describing the empirical strategies we use to identify the effect of surveys on education choice. Particularly, how we exploit the survey sampling scheme to identify the causal effect of being surveyed. This section highlights how – despite successful randomization – we need to deal with non-random sorting across units of randomization.

First, we simply compare the outcomes of those surveyed and those non-surveyed within each municipality. This is given by the linear regression:

$$Y_{ismc} = \delta_0 + \delta_1 Survey_{ismc} + \gamma_m + \varepsilon_{ismc} \quad (2.1)$$

where  $Y_{ismc}$  is the educational outcome for individual  $i$  in school  $s$  in cohort  $c$  and municipality  $m$ ,  $Survey$  is an indicator for whether the individual was surveyed, and  $\gamma_m$  is a municipality fixed effect. Given the survey sampling scheme (cf. Section 2.2.1),  $\delta_1$  can be interpreted as a causal effect of being surveyed if (a) individuals, classes and schools are randomly sampled and (b) there is full compliance as everyone assigned to the survey responded to the survey. In other words, if individuals are neither assigned to sampled classes nor refusing to respond to the survey based on the unobservables,  $\varepsilon_{ismc}$ , in (2.1) which affect the educational outcome of interest.

Second, we use both a within- and a between-school estimator to deal with potential non-random sampling across schools as well as non-random individual non-response. The within-school strategy is ideal if (a') individuals and classes are randomly sampled *within* schools and (b) there is full compliance in student survey response. Random assignment to classes conditional on school (a') is a more credible assumption than (a) if the schools sampled are systematically different from the schools not sampled. The between-school strategy also



assumes (a) random sampling of schools (and classes), but tries to get at potential violations of (b) non-random student non-response by measuring “treatment” at the school level. However, the benefit of not having to assume (b) comes at the cost of the estimated treatment effect being attenuated towards zero.

The following two sub-sections are devoted to providing more details on the within- and between-school estimators we apply. Overall, the empirical strategies trade-off precision and bias in different ways by imposing different identifying assumptions and measuring “survey treatment” at different unit levels. We will discuss the threats to interpreting each of these “survey effects” as causal in even more detail when presenting the data and empirical results in Sections 2.3 and 2.4.

### 2.2.3.1 Within-School

The within-school strategy simply compares the outcomes of individuals who were in the treated classes to those who were in the control classes *within* each school where some students were assigned to treatment. Figure 2.2.2 (a) illustrates this identification strategy. *Municipality A* has three schools of which three classes in *School A* and one class in *School B* are assigned to treatment (marked with light shading). The within-school strategy essentially compares the students in the three treated classes in *School A* to the students in the four control group classes in *School A* and the students in the treated class in *School B* to the students in the remaining three classes in *School B*. In contrast, the within-municipality specification (2.1) simply compares four treated classes in *Municipality A* with the nine control classes in *Municipality A*. This distinction is important if there is significant sorting across schools or if there are influential unobserved school-specific factors affecting educational outcomes. The within-school estimates are given by the linear regression:

$$Y_{ismc} = \beta_0 + \beta_1 Survey_{ismc} + \gamma_s + \varepsilon_{ismc} \quad (2.2)$$

where  $Y_{ismc}$  is the educational outcome for individual  $i$  in school  $s$  in cohort  $c$  and municipality  $m$ ,  $Survey$  is an indicator for being assigned to the survey when in 3rd grade, and  $\gamma_s$  is a school fixed effect.  $\beta_1$  can be interpreted as the causal effect of being surveyed if treated classes are not selected based on unobservables,  $\varepsilon_{ismc}$ , in (2.2) that affect the educational outcome. This seems reasonable based on the class selection criteria outlined in Section 2.2.1. However, there are a few empirical issues we need to deal with. First, we only partially observe class

assignment in 3rd grade as it is only observed for those who are assigned to treatment and comply. Thus we need to impute 3rd grade school assignment by survey response and 9th grade school.<sup>11</sup> Second, there may be attrition due to some students moving after randomization occurred. Third, some students may also have been in a different class, school, municipality, or abroad at the time of randomization but otherwise followed the sampled class. Fourth, many students switch schools between 3rd and 9th grade simply because some schools specialize in either younger or older grades. Fifth, some students (more realistically their parents) might have opted out of the survey for privacy reasons. Therefore, both student non-response and mobility between 3rd and 9th grade pose threats to the identification of  $\beta_1$ . In Section 2.3.2.1, we assess the credibility of the identifying assumptions by testing for balance on a range of variables determined pre-treatment. When possible, we also try to control for location in 3rd grade and whether the student is foreign born in (2.2) – which seems to be a good proxy for mobility.

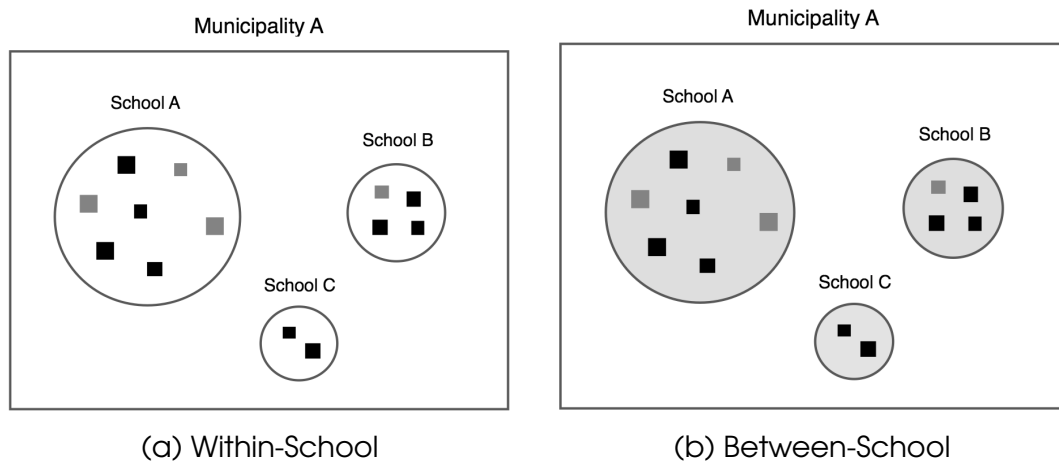


Figure 2.2.2: Within- and Between-School Variation in Treatment Assignment  
 Note: The Figure displays the two types of exogenous variation we exploit in the data. The left panel (a) displays how we compare treated and control (light shading) units within each school when employing the within-school identification strategy. The right panel (b) displays how we compare schools who have some classes assigned to treatment and no classes assigned to treatment, respectively, in the between-school identification strategy.

<sup>11</sup>Section 2.3.1.1 provides more detail on the imputation of schools.

### 2.2.3.2 Between-Schools

To deal with some of the potential threats to identification of a causal survey effect using the simple and the within-school strategy, we also use a between-school empirical strategy. To this end, we compare the outcomes of individuals who were in *ETF schools* (i.e. schools where there was at least one treated class) to those schools who did not have any classes assigned to treatment. Figure 2.2.2 (b) illustrates this identification strategy. *Municipality A* has three schools of which *School A* and *School B* are ETF schools (marked with light shading) and *School C* is a non-ETF school. The between-school strategy essentially compares the average outcomes of the students in the two ETF schools (*School A* and *School B*) to the students in the non-ETF school (*School C*). The between-school estimates are given by the linear regression:

$$Y_{ismc} = \alpha_0 + \alpha_1 ETFschool_{smc} + \gamma_m + \varepsilon_{ismc} \quad (2.3)$$

where  $ETFschool_{smc}$  is an indicator for whether school  $s$  in municipality  $m$  is an ETF school for individuals in cohort  $c$ .  $\alpha_1$  can be interpreted as the causal effect of being in a school where some students were surveyed. In this sense,  $\alpha_1$  can be thought of as an intent to treat effect (ITT). Causal inference naturally rests on the assumption that schools are not sampled based on student unobservables,  $\varepsilon_{ismc}$ , in (2.3).

The main advantage of this between-school estimator is that there is no need to impose assumptions on student survey non-response. There are two main drawbacks. First, as classes were systematically sampled bigger schools are overrepresented within municipality, due to the fact that they have more classes. This problem is particularly severe in small municipalities, where the majority of classes were sampled. This induces a mechanical bias, as small special schools are less likely to be assigned to the treatment. This bias can be corrected controlling for school size or restricting the sample to municipalities where a non-trivial fraction of schools has been left out of the sample (bigger municipalities). Second, some schools – particularly larger schools – have few surveyed students. This means that the estimated “survey effect” will be attenuated towards zero, since the average outcome in “surveyed” ETF schools is an average over a few surveyed students and many non-surveyed students. To accommodate this, we also estimate a specification of (2.3) where we replace  $ETFschool_{smc}$  with the fraction of students in the school who were assigned to the survey,  $FractionSurveyed_{smc}$ . This specification captures the

intensity of treatment in each school, which minimizes the attenuation bias. However,  $FractionSurveyed_{smc}$  may be correlated with the number of students not responding to the survey. To deal with this potential issue, we rather include the predicted fraction of students surveyed in the school,  $\widehat{FractionSurveyed}_{smc}$ , in (2.3). We predict the number of classes sampled in each school by: First, using Maimonides' rule (with a 30 students per class cap) for 9th grade enrollment to predict the average class size.<sup>12</sup> Second, given the number of students surveyed, approximate how many classes were sampled in each school. There were very few cases in which whole classes did not comply to the survey assignment. Thus, this specification of (2.3) should minimize selective non-response bias while also providing a reasonable measure of the intensity of treatment at the school level.

Finally, it should be noted that we assume no spillover effects on untreated units (i.e. individuals, classes, and schools) when making causal inference throughout the paper. That is, we impose the stable unit treatment value assumption (SUTVA) of any unit's outcome being unaffected by another unit's treatment assignment. If this assumption does *not* hold because there are spillovers from those surveyed to their non-surveyed peers in the same school, then both the simple within-municipality ( $\delta_1$ ) and the within-school estimate ( $\beta_1$ ) of the "survey effect" may be attenuated towards zero, while the between-school estimate ( $\alpha_1$ ) will encompass these spillovers. This attenuation bias will presumably be stronger for the within-school estimate ( $\beta_1$ ) than for the within-municipality estimate ( $\delta_1$ ). We can get a sense of the presence and strength of potential spillovers by comparing the different estimates. If there are strong spillovers on peers *within* schools, then the "control classes" in the within-school strategy are partially "treated". This means that the estimated "survey effect" will be attenuated towards zero as the average outcome of those in surveyed classes is compared to the average outcomes of those not surveyed in the same school, but potentially affected by their surveyed peers. For the between-school estimates, these peer spillovers would work in the opposite direction by attenuating the estimated "survey effect" less towards zero, since they compare the average outcomes in schools where some classes are assigned to the survey treatment to the average outcomes in schools where no students are assigned to treatment.

---

<sup>12</sup>Several papers have used Maimonides' rule to estimate the effect of class size on student achievement (Angrist and Lavy, 1999) and long-term outcomes (Fredriksson et al., 2013).

## 2.3 Data

In this section, we describe the data and the sample selection.

We merge the ETF survey data to several administrative registers via the unique Swedish individual identifier. Our measures of educational choices and outcomes originate from several registers administered by Statistics Sweden (*SCB*). We have detailed data on educational choices and outcomes from the 9th grade registry (incl. grades in individual courses), the High School registry (incl. grades in individual courses, grade point average (GPA), track and specialization choices), and the Higher Education registry (incl. detailed educational codes for all enrollment spells, course credits accumulated during enrollment, and acquired degrees).

The Multigeneration registry allows us to link children to their parents. It also contains information on family size and composition. Additional background variables are obtained from the longitudinal integration database for health insurance and labour market studies (*LISA*) from which we have yearly observations during the period 1990-2013. The parental background variables we observe include age, civil status, highest completed education, employment, earnings, and disposable family income. We supplement this with earnings information from the Register Based Labor Market Statistics (*RAMS*) for the years 1986-89 and information on disposable family income from the Income and Tax registry (*IoT*) for the years 1978-89. This means that we can measure disposable family income (parental earnings) from birth (age 3) to age 31 for the youngest cohort and from age 6 (age 14) to 41 for the oldest cohort in our sample.

### 2.3.1 Sample Selection

We focus on the cohorts who completed compulsory schooling (9th grade) in the school-years 1987/88, 1992/93, 1997/98. For children who followed the ordained educational path, this corresponds to the cohorts born in 1972, 1977, and 1982. 95.99%, 95.48% and 5.18% of students ordained to graduate from compulsory school grade 9 in 1988, 1993, 1998 were born in 1972, 1977, and 1982, respectively. For each birth cohort, we also focus exclusively on the sampled municipalities.

For the samples used in the within- and between-school empirical strategies, we further exclude those attending very small schools in 9th grade; i.e. schools

with a graduating cohort of 20 or fewer students. It is difficult to make a reasonable ETF school imputation for these small schools, since we do not know whether a few surveyed students switched to the school or were part of a very small sampled class. This restriction only drops 0.6%, 0.7%, and 1.78% of the sample for the ETF72, ETF77, and ETF82 cohorts, respectively.

### 2.3.1.1 ETF School Imputation

We observe which school everyone attends at the end of compulsory schooling; i.e. in 9th grade. However, we only observe earlier class and school choices for those surveyed. In order to impute school status for the whole sample – which we need for the within and between-school strategies – we need to approximate how each observed 9th grade school corresponds to each 3rd grade school. To this end, we use (class) school codes in (6th) 3rd grade for those assigned to the survey for the (ETF72) ETF77 and ETF82 cohorts. These codes are linked to the school codes we observe for everyone in 9th grade. The information on (class) school codes thus reveals how many surveyed students in the same 3rd grade (class) school are also in the same 9th grade school. The details of the ETF school imputations are as follows:

**ETF72 cohort.** We use the information on classes in grade 6 to approximate whether the grade 9 school code corresponds to an ETF school. Note that most students attend the same school in grades 6 and 9, but most students change school from grade 3 to 9 (or the school code itself changes). We impute ETF schools as follows: (i) For each grade 9 school code, sum the number of students with each grade 6 class code,  $N_s^6$ . (ii) If there are at least two students from a surveyed 6th grade class in the 9th grade school, then we divide  $N_s^6$  by the total number of surveyed students in the 9th grade school. This yields the fraction of students from the same 6th grade class who also attend the same school in 9th grade. A high fraction means that the 9th grade school was a destination school for those attending an ETF school in 3rd grade. We assign ETF school status to the schools for which this fraction was above 0.5. This changes assignment of 30 schools – most of which are very small and drop out of the restricted sample in the between-school analysis; see Sections 2.3.2 and 2.4.

**ETF77 and ETF82 cohorts.** Our imputation procedure is as follows: (i) For each grade 9 school code, find the modal grade 3 school code. (ii) Sum the number of students with the modal school code,  $N_s^3$ . (iii) Divide  $N_s^3$  by the total surveyed students in the 9th grade school,  $s$ . For each grade 9 school code, this

yields the fraction of students with the modal grade 3 school code. (iv) Change school status if fewer than a third of students are coming from the modal grade 3 school. This changes assignment status for one school in the ETF77 cohort and 17 schools in the ETF82 cohort. Note that the number of schools increases over time, but this does not affect school assignments much.

Once control group students are matched up, we calculate the fraction of surveyed students in each grade 9 school. We classify a school to be a non-ETF school if: there are five or fewer surveyed students in the grade 9 school *and* they represent at most 5% of the school. This changes status for 3, 62, and 35 schools in the ETF72, ETF77, and ETF82 cohort, respectively.<sup>13</sup>

### 2.3.2 Descriptive Statistics

Table 2.3.1 shows that survey response was high – initially above 90% – but there was also some attrition as response rates fell by about 20 percentage points over the seven year period from 3rd to 10th grade.

Table 2.3.1: Percentage of students and parents responding to the survey

	Child Survey			Parent Survey	
	Grade 3	Grade 6	Grade 10	Grade 3	Grade 6
Cohort 1972	93	85	72	75	–
Cohort 1977	95	91	73	–	77
Cohort 1982	–	87	–	–	75

Table 2.3.2 displays the number of students and schools in the sampled municipalities. Around 80% of schools were sampled in the 72 and 82 cohorts, while only 55% of schools were sampled for the 77 cohort. The fraction of students sampled within each sampled school is also highest (38%) in the 72 cohort and lowest (26%) in the 77 cohort. Overall, the table shows that there should be enough variation in survey assignment within-municipalities to estimate (2.1) and within-schools to estimate (2.2).

<sup>13</sup>Future versions of the paper will provide sensitivity analysis of the importance of the chosen thresholds. We also intend to use Maimonides' rule in order to exclude only municipalities who had almost all classes sampled in expectation.

To assure that there is enough variation between schools for the between-school estimates of (2.3), we also conduct the analysis on a restricted sample of municipalities where at most 85% of students are assigned to the survey treatment. This retains mostly big municipalities; see Section 2.2.1 for details. Table 2.3.3 presents the equivalent variation to Table 2.3.2 for this restricted sample. Naturally, both the fraction of sampled schools within the municipality and the fraction of sampled students within each sampled school are lower in this sample of larger municipalities and schools.

Table 2.3.2: Students and schools sampled - Unrestricted sample

	Students	Sampled Schools	Schools	Fraction in Sampled Schools	Fraction Surveyed in Sampled Schools
Cohort 1972	23305	143	184	0.82	38.2
Cohort 1977	24692	120	235	0.55	25.8
Cohort 1982	28709	225	290	0.79	28.1

Table 2.3.3: Students and schools sampled - Restricted sample

	Students	Sampled Schools	Schools	Fraction in Sampled Schools	Fraction Surveyed in Sampled Schools
Cohort 1972	9232	39	79	0.56	15.8
Cohort 1977	16198	45	155	0.33	15.7
Cohort 1982	19883	143	204	0.71	24.4

Sample restricted to municipalities where at least 85% of the students are not part of an EFT school

### 2.3.2.1 Balancing Tests

To corroborate the randomness of the sampling scheme, we perform a number of balancing tests on the pre-determined characteristics of the “treatment” and “control” groups.

Tables 2.A.1 to 2.A.13 in Appendix 2.A display the balancing tests for each of the cohorts, under different sample restrictions, and for each of the four empirical strategies. Each table displays control group means in the first column. We also present three sets of balancing tests. First, regression tests



without and with controls. Second, standardized difference, the difference between the treatment group mean and the control group mean of each observed characteristic,  $X$ , scaled by the pooled variance.

The balancing tests indicate that assumption (b) in Section 2.2.3 of random survey non-response is more of an empirical issue – especially for the 72 cohort and to some extent for the 82 cohort – as there are some systematic and significant differences in observed characteristics between “treatment” and “control” groups for the within-school strategy that are not present for the between-school strategy. Particularly, having divorced parents and being foreign born are predictive of group assignment. Students with disrupted families might have been more likely to drop out of the survey. Being foreign born presumably introduces issues of selective mobility affecting the within-school “survey effect” estimate. To get a sense of how important selective mobility may be, we also control for location in 3rd grade (when possible) and whether the student is foreign born in (2.2) – which seems to be a good proxy for mobility.<sup>14</sup> 92%, 93%, and 95% of the students in the 72, 77, and 82 cohorts, respectively, remain in one of the sampled municipalities. Thus mobility should not have a major impact on the analysis. Nevertheless, we also perform the analysis on a restricted sample of Swedish born students only. This improves balance on pre-determined characteristics, particularly for the 1977 and 1982 cohort.

All in all it appears that different identification strategies work for different cohorts. In particular, when restricting the sample to Swedish born students, there are no statistically significant differences between surveyed and non-surveyed students within municipality for the 1977 and 1982 cohorts (see Tables 2.A.6 and 2.A.11). Controlling for school size and restricting the sample to the bigger municipalities, there are no statistically significant differences between students in sampled and non-sampled schools for the 1972 and 1982 cohorts (see Tables 2.A.4 and 2.A.13).

---

<sup>14</sup>Whether parents divorced is also related to student mobility and may even be related to student educational outcomes. In future versions of the paper, we plan to explore the importance of this and other mobility channels further as we have data on the timing of parental divorce and the timing of migration – in and out of Sweden as well as between Swedish municipalities.

## 2.4 Empirical Results

This section presents the empirical results. We observe complete educational spells from the last year of compulsory schooling (9th grade) as well as individual employment status and earnings throughout the early careers. We focus on shorter- and longer-term outcomes. The short-term outcomes are 9th grade GPA, and indicators for whether advanced Math and English were chosen in 9th grade for the 72 and 77 cohort, while the short-term outcomes for the 82 cohort are the individual grades in 9th grade Math, English, and Swedish.<sup>15</sup> The long-term outcomes include highest completed educational attainment as well as average earnings and days unemployed during the year when the individuals are 28-31. Tables 2.B.1 to 2.B.12 in Appendix 2.B display the empirical results for each of the cohorts, and each of the four empirical strategies. Each table presents estimates of the “survey effect” on a short- and a long-term index in order to assess whether the survey had *any* effect on educational choices and subsequent outcomes. These indexes are constructed by standardizing each variable that enters the index to the control mean and standard deviation, and then taking the average of the standardized variables with signs such that larger is better. For the within-school strategy, indexes are constructed at the school level; i.e. the control group is the students who have *not* been sampled in each school.<sup>16</sup>

We focus for each cohort on the specifications that pass the robustness tests, and conclude that there is overall no effect of the survey on short and long-term effects, for both students with low and highly educated families. We find no effect of the treatment on the main indexes and their components for the 1972 cohort (between-school specification, Tables 2.B.3 and 2.B.4), and for the 1982 cohort (within-municipality specification, Table 2.B.9). However we do find some reduction in unemployment days for surveyed students from low-education families for the 1972 cohort (within-municipality specification, Table 2.B.5), and an increase in the graduation rate from short-college for the 1982 cohort for students with educated parents (between-school specification, Table

---

<sup>15</sup>The short-term outcomes differ across cohorts because of institutional changes in schools and data availability. Bjorklund et al. (2005) provide more details on these institutional changes that also drive the increased number of schools for the youngest cohort.

<sup>16</sup>This way of constructing outcome indexes builds on Kling et al. (2007).

2.B.11). The latter result is not consistent with what we find with the within-municipality specification, but the two samples are not directly comparable due to the different restrictions.

#### 2.4.1 Future directions

To understand whether, beyond the general zero effect, there actually is an heterogeneous effect of being surveyed, we need to consider the problem of multiple hypothesis testing. In the next step of the analysis we will use the testing procedure outlined in List et al. (2016), which builds on the stepwise procedures developed in Romano and Wolf (2005a,b, 2010) and was extended to heterogeneous treatment effects by Lee and Shaikh (2014) in the context of the PROGRESA program.<sup>17</sup> This testing procedure asymptotically controls for the familywise error rate (i.e. the probability of one false rejection) and is asymptotically balanced such that all marginal probabilities of rejecting any true null hypothesis are approximately equal. This testing procedure has better power properties as it incorporates information on the joint dependence structure of the test statistics when determining which null hypotheses to reject. Thus at a significance level of  $\alpha$ , all rejected null hypothesis are actually false with probability  $1 - \alpha$ . We will use this procedure to adjust the standard errors when making inference for multiple outcomes (three short- and six long-term) and multiple subgroups (low and high parental education).

Secondly, we might be finding a zero effect of being surveyed for the simple reason that our surveys were assigned *after* students had taken important choices (elective courses, high school enrollment and track choice). In future versions of this paper, we will focus on whether the parent survey had an impact on the education choices of non-surveyed siblings. We will present separate estimates for siblings who were ordained to attend grades  $g = 10, 11, 12$  (17-19 years old) and thus high school bound. These siblings are at a critical junction on their

---

<sup>17</sup>Chapter 15 in Lehmann and Romano (2006) provides an overview of these testing procedures. Similar adjustments for multiple testing have been made in recent analysis of the HighScope Perry Preschool Program (Heckman et al., 2010, 2011, 2013), Anderson (2008) who analyzes early childhood interventions, and Kling et al. (2007) who analyze the Moving to Opportunity (MTO) experiment when drawing inference about the effect of the program on multiple outcomes using closely related results on stepwise multiple testing developed in Westfall and Young (1993).

schooling trajectory. Thus they may be particularly sensitive to the influence of their parents becoming more aware of the importance of schooling.

Thirdly, to investigate potential spillover effects we will provide separate estimates for non-surveyed siblings who are attending compulsory schooling grades  $g = 1, 2, \dots, 9$  (7-16 years old) when their parent was administered a survey (at random) because they had a sibling in one of the ETF cohorts.

## 2.5 Conclusion

Most empirical advances in the social and health sciences over the past decades have depended crucially on the use of survey data. If surveying individuals draws their attention to risks, returns, or choices previously not salient to them and this changes their subsequent behavior, then it may bias parameter estimates and conclusions drawn from survey data.

We assess whether surveys causally changed educational choices and outcomes of students attending compulsory school in Sweden in the 80s and early 90s. We do not find strong reasons to worry about extensive surveying changing educational choices and subsequent outcomes. This is reassuring for both the external and internal validity of estimates based on (this) survey data.

There are, however, some cases for which the survey increased educational attainment and job stability of the surveyed individuals. We thus think it is worth investigating in future research, after properly correcting for multiple hypothesis testing, whether being surveyed affect choices for specific subpopulations. To further get at mechanisms, we will also extend our analysis to the siblings of the surveyed students, for whom parental surveys might have revealed information at critical junctions on their schooling trajectory.

## 2.A Balancing Tests

### 2.A.1 1972 Cohort

Table 2.A.1:  
Differences in student background by treatment status:  
Within-school and municipality specifications  
1972 cohort - All students

Variable	Control Mean	Difference(Sch)	Difference(Munic)	Cohen's d
Female	0.49 (0.50)	-0.00 (0.01)	-0.00 (0.01)	0.00
Foreign born	0.07 (0.26)	-0.03*** (0.00)	-0.02*** (0.00)	0.19
Swedish born, foreign parent	0.11 (0.32)	0.00 (0.01)	0.00 (0.01)	0.05
Divorced parent	0.25 (0.43)	-0.03*** (0.01)	-0.03*** (0.01)	0.18
Father's disp. income (1000 kr, age 6-9)	153.03 (51.79)	2.20 (1.46)	2.84 (2.05)	0.06
Mother's disp. income (1000 kr, age 6-9)	94.26 (49.71)	-1.08 (1.02)	-0.87 (1.11)	0.25
Father with low SES	0.36 (0.48)	0.02* (0.01)	0.01 (0.01)	-0.12
Father with medium SES	0.42 (0.49)	0.00 (0.01)	0.00 (0.01)	0.01
Father with high SES	0.23 (0.42)	-0.02** (0.01)	-0.01 (0.02)	0.12
Father educ: high school	0.44 (0.50)	-0.02 (0.01)	-0.02* (0.01)	0.03
Father educ: college or more	0.26 (0.44)	-0.02* (0.01)	-0.01 (0.01)	0.14
Father: in the labor force	0.92 (0.27)	0.01*** (0.01)	0.01** (0.00)	-0.10
Mother with low SES	0.37 (0.48)	0.00 (0.01)	-0.00 (0.01)	-0.17
Mother with medium SES	0.51 (0.50)	0.01 (0.01)	0.01 (0.01)	0.09
Mother with high SES	0.12 (0.33)	-0.01 (0.01)	-0.00 (0.01)	0.12
Mother educ: high school	0.42 (0.49)	0.00 (0.01)	0.01 (0.01)	-0.07
Mother educ: college or more	0.29 (0.45)	-0.01 (0.01)	-0.01 (0.02)	0.14
Mother: in the labor force	0.92 (0.26)	0.01*** (0.01)	0.01*** (0.00)	-0.06
Hotelling's T-squared	413.69			

\*  $p < 0.10$ , \*\*  $p < 0.05$ , \*\*\*  $p < 0.01$

Column 2 and 3 test for differences respectively within school and municipality.  
No sample restriction. Controls: none. Prices adjusted to 2014: 1 SEK = 0.11 €. Standard errors clustered at the school level.

Table 2.A.2:  
Differences in student background by treatment status:  
Within-school and municipality specifications  
1972 cohort - Swedish born

Variable	Control Mean	Difference(Sch)	Difference(Munic)	Cohen's d
Female	0.49 (0.50)	-0.00 (0.01)	-0.00 (0.01)	0.00
Foreign born	0.00 (0.00)	0.00	0.00	0.19
Swedish born, foreign parent	0.00 (0.00)	0.00	0.00	0.05
Divorced parent	0.24 (0.43)	-0.03*** (0.01)	-0.03*** (0.01)	0.18
Father's disp. income (1000 kr, age 6-9)	157.13 (51.56)	2.20 (1.51)	2.62 (1.91)	0.06
Mother's disp. income (1000 kr, age 6-9)	92.17 (50.20)	-1.97* (1.05)	-1.73 (1.17)	0.25
Father with low SES	0.32 (0.47)	0.02** (0.01)	0.01 (0.01)	-0.12
Father with medium SES	0.43 (0.50)	-0.00 (0.01)	-0.00 (0.01)	0.01
Father with high SES	0.24 (0.43)	-0.02* (0.01)	-0.01 (0.02)	0.12
Father educ: high school	0.44 (0.50)	-0.01 (0.01)	-0.02 (0.01)	0.03
Father educ: college or more	0.28 (0.45)	-0.03** (0.01)	-0.01 (0.02)	0.14
Father: in the labor force	0.95 (0.22)	0.01** (0.00)	0.01* (0.00)	-0.10
Mother with low SES	0.33 (0.47)	0.01 (0.01)	0.00 (0.01)	-0.17
Mother with medium SES	0.54 (0.50)	0.00 (0.01)	0.00 (0.01)	0.09
Mother with high SES	0.13 (0.33)	-0.01 (0.01)	-0.00 (0.01)	0.12
Mother educ: high school	0.43 (0.50)	0.01 (0.01)	0.01 (0.01)	-0.07
Mother educ: college or more	0.30 (0.46)	-0.01 (0.01)	-0.01 (0.02)	0.14
Mother: in the labor force	0.95 (0.22)	0.01* (0.01)	0.01 (0.01)	-0.06
Hotelling's T-squared	339.22			

\*  $p < 0.10$ , \*\*  $p < 0.05$ , \*\*\*  $p < 0.01$

Column 2 and 3 test for differences respectively within school and municipality. Sample restricted to Swedish born students. Controls: none. Prices adjusted to 2014: 1 SEK = 0.11 €. Standard errors clustered at the school level.

Table 2.A.3:  
Differences in student background by treatment assignment:  
Between-school specification  
1972 cohort - All students

Variable	Control Mean	No controls Difference	Controls Difference	Cohen's d
Female	0.49 (0.50)	-0.01 (0.01)	-0.01* (0.01)	0.01
Foreign born	0.09 (0.28)	-0.03* (0.02)	-0.02 (0.02)	0.11
Swedish born, foreign parent	0.13 (0.34)	-0.03*** (0.01)	-0.03*** (0.01)	0.10
Divorced parent	0.28 (0.45)	-0.06*** (0.01)	-0.05*** (0.01)	0.15
Father's disp. income (1000 kr, age 6-9)	153.79 (55.36)	-2.69 (2.83)	-4.51 (3.09)	0.05
Mother's disp. income (1000 kr, age 6-9)	101.41 (52.21)	-14.18*** (2.52)	-12.54*** (2.66)	0.29
Father with low SES	0.37 (0.48)	0.01 (0.02)	0.03 (0.03)	-0.03
Father with medium SES	0.42 (0.49)	-0.01 (0.01)	-0.01 (0.01)	0.01
Father with high SES	0.21 (0.41)	-0.01 (0.02)	-0.02 (0.02)	0.02
Father educ: high school	0.44 (0.50)	-0.01 (0.01)	-0.01 (0.01)	0.02
Father educ: college or more	0.25 (0.43)	-0.01 (0.02)	-0.03 (0.03)	0.03
Father: in the labor force	0.92 (0.28)	0.02* (0.01)	0.01 (0.01)	-0.06
Mother with low SES	0.34 (0.47)	0.08*** (0.03)	0.09*** (0.03)	-0.16
Mother with medium SES	0.54 (0.50)	-0.06*** (0.02)	-0.07*** (0.02)	0.12
Mother with high SES	0.12 (0.33)	-0.02 (0.01)	-0.02 (0.01)	0.05
Mother educ: high school	0.42 (0.49)	0.01 (0.01)	0.01 (0.01)	-0.02
Mother educ: college or more	0.29 (0.45)	-0.02 (0.02)	-0.04 (0.02)	0.06
Mother: in the labor force	0.92 (0.27)	0.01 (0.01)	0.00 (0.01)	-0.03
Hotelling's T-squared	291.91			

\*  $p < 0.10$ , \*\*  $p < 0.05$ , \*\*\*  $p < 0.01$

No sample restriction. The specification with controls includes a control for school size. Prices adjusted to 2014: 1 SEK = 0.11 €. Standard errors clustered at the school level.

Table 2.A.4:  
Differences in student background by treatment assignment:  
Between-school specification  
1972 cohort - Restricted sample

Variable	Control Mean	No controls Difference	Controls Difference	Cohen's d
Female	0.49 (0.50)	-0.00 (0.01)	-0.01 (0.01)	0.01
Foreign born	0.09 (0.28)	0.01 (0.02)	0.01 (0.02)	0.11
Swedish born, foreign parent	0.13 (0.34)	-0.01 (0.01)	-0.01 (0.01)	0.10
Divorced parent	0.28 (0.45)	-0.01 (0.02)	0.01 (0.02)	0.15
Father's disp. income (1000 kr, age 6-9)	153.67 (55.88)	5.70 (4.12)	2.04 (4.81)	0.05
Mother's disp. income (1000 kr, age 6-9)	101.85 (52.19)	-3.59 (3.28)	-0.62 (3.37)	0.29
Father with low SES	0.37 (0.48)	-0.05 (0.04)	-0.02 (0.04)	-0.03
Father with medium SES	0.42 (0.49)	-0.01 (0.02)	-0.00 (0.02)	0.01
Father with high SES	0.21 (0.41)	0.06* (0.03)	0.03 (0.04)	0.02
Father educ: high school	0.44 (0.50)	-0.03 (0.02)	-0.02 (0.02)	0.02
Father educ: college or more	0.25 (0.43)	0.08** (0.04)	0.04 (0.04)	0.03
Father: in the labor force	0.91 (0.28)	-0.00 (0.01)	-0.01 (0.01)	-0.06
Mother with low SES	0.34 (0.47)	-0.02 (0.04)	-0.01 (0.04)	-0.16
Mother with medium SES	0.54 (0.50)	-0.02 (0.03)	-0.02 (0.03)	0.12
Mother with high SES	0.12 (0.33)	0.04* (0.02)	0.03 (0.02)	0.05
Mother educ: high school	0.42 (0.49)	-0.04** (0.02)	-0.02 (0.02)	-0.02
Mother educ: college or more	0.29 (0.45)	0.06 (0.04)	0.03 (0.04)	0.06
Mother: in the labor force	0.92 (0.27)	-0.00 (0.01)	-0.01 (0.02)	-0.03
Hotelling's T-squared	100.44			

\*  $p < 0.10$ , \*\*  $p < 0.05$ , \*\*\*  $p < 0.01$

Sample restricted to municipalities where less than 85% of the students are in sampled schools. The specification with controls includes a control for school size. Prices adjusted to 2014: 1 SEK = 0.11 €. Standard errors clustered at the school level.



## 2.A.2 1977 Cohort

Table 2.A.5:  
Differences in student background by treatment status:  
Within-school and municipality specifications  
1977 cohort - All students

Variable	Control Mean	Difference(Sch)	Difference(Munic)	Cohen's d
Female	0.49 (0.50)	0.01 (0.01)	0.01 (0.01)	-0.01
Foreign born	0.09 (0.28)	-0.02*** (0.01)	-0.02*** (0.00)	0.18
Swedish born, foreign parent	0.10 (0.30)	0.00 (0.01)	-0.00 (0.01)	0.06
Divorced parent	0.18 (0.38)	-0.02** (0.01)	-0.02* (0.01)	0.11
Father Disp. income (1000 kr, age 6-9)	179.96 (83.14)	-1.65 (1.66)	-1.13 (1.91)	0.06
Mother Disp. income (1000 kr, age 6-9)	129.60 (50.87)	-0.08 (1.11)	0.41 (1.10)	0.12
Father with low SES	0.37 (0.48)	0.01 (0.01)	-0.00 (0.02)	-0.08
Father with medium SES	0.40 (0.49)	0.01 (0.01)	0.01 (0.01)	-0.02
Father with high SES	0.23 (0.42)	-0.02 (0.01)	-0.01 (0.01)	0.11
Father educ: high school	0.43 (0.50)	0.01 (0.01)	0.00 (0.01)	-0.02
Father educ: college or more	0.29 (0.45)	-0.01 (0.01)	-0.00 (0.01)	0.10
Father: In labor force (1990)	0.94 (0.24)	0.01* (0.00)	0.01** (0.00)	-0.10
Mother with low SES	0.38 (0.49)	0.03** (0.01)	0.01 (0.01)	-0.11
Mother with medium SES	0.51 (0.50)	-0.02 (0.01)	-0.01 (0.01)	0.05
Mother with high SES	0.12 (0.32)	-0.01 (0.01)	-0.00 (0.01)	0.08
Mother educ: high school	0.44 (0.50)	0.01 (0.01)	-0.00 (0.01)	-0.06
Mother educ: college or more	0.30 (0.46)	-0.01 (0.01)	-0.00 (0.01)	0.08
Mother: In labor force (1990)	0.93 (0.26)	0.01* (0.00)	0.01 (0.01)	-0.08
Hotelling's T-squared	95.78			

\*  $p < 0.10$ , \*\*  $p < 0.05$ , \*\*\*  $p < 0.01$

Column 2 and 3 test for differences respectively within school and municipality. No sample restriction. Controls: none. Prices adjusted to 2014: 1 SEK = 0.11€. Standard errors clustered at the school level.

Table 2.A.6:  
Differences in student background by treatment status:  
Within-school and municipality specifications  
1977 cohort - Swedish born

Variable	Control Mean	Difference(Sch)	Difference(Munic)	Cohen's d
Female	0.48 (0.50)	0.02 (0.01)	0.01 (0.01)	-0.01
Foreign born	0.00 (0.00)	0.00	0.00	0.18
Swedish born, foreign parent	0.00 (0.00)	0.00	0.00	0.06
Divorced parent	0.16 (0.37)	-0.02** (0.01)	-0.01 (0.01)	0.11
Father Disp. income (1000 kr, age 6-9)	186.94 (84.36)	-2.23 (1.58)	-1.66 (1.77)	0.06
Mother Disp. income (1000 kr, age 6-9)	127.47 (49.09)	-0.28 (1.26)	0.24 (1.24)	0.12
Father with low SES	0.34 (0.47)	0.00 (0.01)	-0.00 (0.02)	-0.08
Father with medium SES	0.42 (0.49)	0.01 (0.01)	0.01 (0.01)	-0.02
Father with high SES	0.24 (0.43)	-0.02 (0.01)	-0.01 (0.01)	0.11
Father educ: high school	0.43 (0.50)	0.01 (0.01)	0.01 (0.01)	-0.02
Father educ: college or more	0.30 (0.46)	-0.01 (0.01)	-0.00 (0.01)	0.10
Father: In labor force (1990)	0.97 (0.18)	0.00 (0.00)	0.00 (0.00)	-0.10
Mother with low SES	0.34 (0.48)	0.02 (0.01)	0.01 (0.01)	-0.11
Mother with medium SES	0.54 (0.50)	-0.02 (0.01)	-0.01 (0.01)	0.05
Mother with high SES	0.12 (0.32)	-0.00 (0.01)	0.00 (0.01)	0.08
Mother educ: high school	0.45 (0.50)	0.01 (0.01)	-0.00 (0.01)	-0.06
Mother educ: college or more	0.32 (0.47)	-0.01 (0.01)	0.00 (0.01)	0.08
Mother: In labor force (1990)	0.96 (0.21)	0.01 (0.01)	0.01 (0.00)	-0.08
Hotelling's T-squared	89.16			

\*  $p < 0.10$ , \*\*  $p < 0.05$ , \*\*\*  $p < 0.01$

Column 2 and 3 test for differences respectively within school and municipality. Sample restricted to Swedish born students. Controls: none. Prices adjusted to 2014: 1 SEK = 0.11 €. Standard errors clustered at the school level.

Table 2.A.7:  
Differences in student background by treatment assignment:  
Between-school specification  
1977 cohort - All students

Variable	Control Mean	No controls Difference	Controls Difference	Cohen's d
Female	0.49 (0.50)	-0.01 (0.01)	-0.01 (0.01)	0.01
Foreign born	0.11 (0.31)	-0.03** (0.01)	-0.03** (0.01)	0.10
Swedish born, foreign parent	0.12 (0.32)	-0.03*** (0.01)	-0.03*** (0.01)	0.09
Divorced parent	0.19 (0.39)	-0.03*** (0.01)	-0.03*** (0.01)	0.08
Father Disp. income (1000 kr, age 6-9)	180.21 (92.25)	-2.49 (3.76)	-4.08 (3.80)	0.03
Mother Disp. income (1000 kr, age 6-9)	133.34 (51.63)	-7.96*** (1.84)	-8.13*** (1.85)	0.16
Father with low SES	0.36 (0.48)	0.05** (0.02)	0.05*** (0.02)	-0.09
Father with medium SES	0.40 (0.49)	-0.00 (0.01)	-0.01 (0.01)	0.00
Father with high SES	0.24 (0.43)	-0.04*** (0.02)	-0.05*** (0.02)	0.11
Father educ: high school	0.43 (0.50)	0.00 (0.01)	0.00 (0.01)	-0.00
Father educ: college or more	0.31 (0.46)	-0.04** (0.02)	-0.05*** (0.02)	0.10
Father: Time in labor force	0.93 (0.26)	0.02*** (0.01)	0.02*** (0.01)	-0.08
Mother with low SES	0.36 (0.48)	0.05** (0.02)	0.05*** (0.02)	-0.10
Mother with medium SES	0.51 (0.50)	-0.02* (0.01)	-0.03* (0.01)	0.05
Mother with high SES	0.13 (0.33)	-0.02** (0.01)	-0.03*** (0.01)	0.07
Mother educ: high school	0.42 (0.49)	0.03** (0.01)	0.03** (0.01)	-0.05
Mother educ: college or more	0.32 (0.47)	-0.04** (0.02)	-0.05*** (0.02)	0.10
Mother: Time in labor force	0.92 (0.27)	0.01* (0.01)	0.01* (0.01)	-0.06
Hotelling's T-squared	228.37			

\*  $p < 0.10$ , \*\*  $p < 0.05$ , \*\*\*  $p < 0.01$

No sample restriction. The specification with controls includes a control for school size. Prices adjusted to 2014: 1 SEK = 0.11 €. Standard errors clustered at the school level.

Table 2.A.8:  
Differences in student background by treatment assignment:  
Between-school specification  
1977 cohort - Restricted sample

Variable	Control Mean	No controls Difference	Controls Difference	Cohen's d
Female	0.49 (0.50)	-0.01 (0.01)	-0.01 (0.01)	0.01
Foreign born	0.11 (0.32)	-0.00 (0.02)	0.00 (0.02)	0.10
Swedish born, foreign parent	0.12 (0.32)	-0.02** (0.01)	-0.02* (0.01)	0.09
Divorced parent	0.19 (0.40)	-0.01 (0.01)	-0.01 (0.01)	0.08
Father Disp. income (1000 kr, age 6-9)	180.58 (93.15)	-2.19 (5.27)	-5.40 (5.19)	0.03
Mother Disp. income (1000 kr, age 6-9)	133.90 (51.74)	-4.97** (2.48)	-4.47* (2.52)	0.16
Father with low SES	0.35 (0.48)	0.04 (0.03)	0.06** (0.03)	-0.09
Father with medium SES	0.40 (0.49)	-0.01 (0.01)	-0.02 (0.01)	0.00
Father with high SES	0.25 (0.43)	-0.03 (0.02)	-0.04* (0.02)	0.11
Father educ: high school	0.43 (0.50)	-0.01 (0.01)	-0.01 (0.01)	-0.00
Father educ: college or more	0.31 (0.46)	-0.03 (0.03)	-0.04 (0.03)	0.10
Father: Time in labor force	0.93 (0.26)	0.01 (0.01)	0.00 (0.01)	-0.08
Mother with low SES	0.36 (0.48)	0.03 (0.03)	0.04 (0.03)	-0.10
Mother with medium SES	0.51 (0.50)	-0.02 (0.02)	-0.02 (0.02)	0.05
Mother with high SES	0.13 (0.33)	-0.01 (0.02)	-0.02 (0.02)	0.07
Mother educ: high school	0.42 (0.49)	0.01 (0.02)	0.02 (0.02)	-0.05
Mother educ: college or more	0.32 (0.47)	-0.02 (0.03)	-0.04 (0.03)	0.10
Mother: Time in labor force	0.92 (0.28)	0.01 (0.01)	0.00 (0.01)	-0.06
Hotelling's T-squared	78.32			

\*  $p < 0.10$ , \*\*  $p < 0.05$ , \*\*\*  $p < 0.01$

Sample restricted to municipalities where less than 85% of the students are in sampled schools. The specification with controls includes a control for school size. Prices adjusted to 2014: 1 SEK = 0.11 €. Standard errors clustered at the school level.

## 2.A.3 1982 Cohort

Table 2.A.9:  
Differences in student background by treatment status:  
Within-school and municipality specifications  
1982 cohort - All students

Variable	Control Mean	Difference(Sch)	Difference(Munic)	Cohen's d
Female	0.50 (0.50)	-0.00 (0.01)	-0.00 (0.01)	0.01
Foreign born	0.13 (0.34)	-0.04*** (0.01)	-0.05*** (0.01)	0.21
Swedish born, foreign parent	0.11 (0.31)	0.02*** (0.01)	0.01** (0.01)	-0.01
Divorced parent	0.14 (0.34)	-0.01* (0.01)	-0.01** (0.01)	0.07
Father's Disp. income (1000 kr, age 6 to 9)	206.55 (146.21)	8.06** (3.87)	6.52** (3.15)	0.01
Mother's Disp. income (1000 kr, age 6 to 9)	162.47 (73.93)	-0.41 (1.00)	-1.17 (1.29)	0.09
Father with low SES	0.36 (0.48)	-0.02** (0.01)	-0.01 (0.01)	-0.08
Father with medium SES	0.39 (0.49)	0.00 (0.01)	0.00 (0.01)	0.01
Father with high SES	0.25 (0.43)	0.02** (0.01)	0.01 (0.01)	0.09
Father educ: high school	0.41 (0.49)	0.00 (0.01)	0.01 (0.01)	-0.06
Father educ: college or more	0.36 (0.48)	0.02* (0.01)	0.01 (0.01)	0.10
Father: In labor force (1991)	0.92 (0.27)	0.02*** (0.00)	0.02*** (0.00)	-0.10
Mother with low SES	0.37 (0.48)	-0.00 (0.01)	0.01 (0.01)	-0.12
Mother with medium SES	0.49 (0.50)	0.00 (0.01)	-0.00 (0.01)	0.05
Mother with high SES	0.14 (0.34)	-0.00 (0.01)	-0.00 (0.00)	0.10
Mother educ: high school	0.42 (0.49)	0.01* (0.01)	0.02** (0.01)	-0.09
Mother educ: college or more	0.37 (0.48)	0.01 (0.01)	0.01 (0.01)	0.09
Mother: In labor force (1991)	0.90 (0.30)	0.01*** (0.00)	0.02** (0.01)	-0.06
Hotelling's T-squared	145.03			

\*  $p < 0.10$ , \*\*  $p < 0.05$ , \*\*\*  $p < 0.01$

Column 2 and 3 test for differences respectively within school and municipality. No sample restriction. Controls: none. Prices adjusted to 2014: 1 SEK = 0.11 €. Standard errors clustered at the school level.

Table 2.A.10: Differences in student background by treatment status:  
Within-school and municipality specifications  
1982 cohort - Same municipality in grades 3 and 9

Variable	Control Mean	Difference(Sch)	Difference(Munic)	Cohen's d
Female	0.49 (0.50)	-0.00 (0.01)	-0.00 (0.01)	0.01
Foreign born	0.08 (0.28)	-0.01** (0.00)	-0.01* (0.01)	0.21
Swedish born, foreign parent	0.11 (0.31)	0.01** (0.01)	0.01* (0.01)	-0.01
Divorced parent	0.13 (0.34)	-0.01 (0.01)	-0.01 (0.01)	0.07
Father's Disp. income (1000 kr, age 6 to 9)	207.32 (147.56)	7.51* (4.05)	6.11* (3.49)	0.01
Mother's Disp. income (1000 kr, age 6 to 9)	162.80 (74.41)	-0.46 (1.02)	-1.26 (1.34)	0.09
Father with low SES	0.36 (0.48)	-0.02* (0.01)	-0.01 (0.01)	-0.08
Father with medium SES	0.40 (0.49)	-0.00 (0.01)	-0.00 (0.01)	0.01
Father with high SES	0.24 (0.43)	0.02** (0.01)	0.01* (0.01)	0.09
Father educ: high school	0.41 (0.49)	0.01 (0.01)	0.01 (0.01)	-0.06
Father educ: college or more	0.37 (0.48)	0.01 (0.01)	0.01 (0.01)	0.10
Father: In labor force (1991)	0.93 (0.26)	0.01*** (0.00)	0.01*** (0.00)	-0.10
Mother with low SES	0.37 (0.48)	0.00 (0.01)	0.01 (0.01)	-0.12
Mother with medium SES	0.49 (0.50)	-0.00 (0.01)	-0.00 (0.01)	0.05
Mother with high SES	0.14 (0.34)	-0.00 (0.01)	-0.00 (0.00)	0.10
Mother educ: high school	0.42 (0.49)	0.02* (0.01)	0.02** (0.01)	-0.09
Mother educ: college or more	0.37 (0.48)	0.01 (0.01)	0.00 (0.01)	0.09
Mother: In labor force (1991)	0.91 (0.29)	0.01* (0.00)	0.01* (0.00)	-0.06
Hotelling's T-squared	123.88			

\*  $p < 0.10$ , \*\*  $p < 0.05$ , \*\*\*  $p < 0.01$

Column 2 and 3 test for differences respectively within school and municipality. Sample restricted to students living in the same municipality in grade 3 and 9. Controls: none. Prices adjusted to 2014: 1 SEK = 0.11 €. Standard errors clustered at the school level.

Table 2.A.11:  
Differences in student background by treatment status:  
Within-school and municipality specifications  
1982 cohort - Swedish born, same municipality in grades 3 and 9

Variable	Control Mean	Difference(Sch)	Difference(Munic)	Cohen's d
Female	0.49 (0.50)	-0.00 (0.01)	0.00 (0.01)	0.01
Foreign born	0.00 (0.00)	0.00	0.00	0.21
Swedish born, foreign parent	0.00 (0.00)	0.00	0.00	-0.01
Divorced parent	0.11 (0.31)	-0.01 (0.01)	-0.00 (0.01)	0.07
Father's Disp. income (1000 kr, age 6 to 9)	221.58 (161.91)	7.48 (4.90)	4.80 (4.10)	0.01
Mother's Disp. income (1000 kr, age 6 to 9)	164.03 (63.93)	-1.14 (1.12)	-2.04 (1.70)	0.09
Father with low SES	0.33 (0.47)	-0.01 (0.01)	-0.00 (0.01)	-0.08
Father with medium SES	0.41 (0.49)	-0.01 (0.01)	-0.01 (0.01)	0.01
Father with high SES	0.26 (0.44)	0.02* (0.01)	0.01 (0.01)	0.09
Father educ: high school	0.41 (0.49)	0.01 (0.01)	0.01 (0.01)	-0.06
Father educ: college or more	0.39 (0.49)	0.01 (0.01)	-0.01 (0.01)	0.10
Father: In labor force (1991)	0.97 (0.18)	0.01* (0.00)	0.00 (0.00)	-0.10
Mother with low SES	0.34 (0.47)	0.01 (0.01)	0.02 (0.01)	-0.12
Mother with medium SES	0.52 (0.50)	-0.01 (0.01)	-0.01 (0.01)	0.05
Mother with high SES	0.14 (0.35)	-0.00 (0.01)	-0.01 (0.01)	0.10
Mother educ: high school	0.43 (0.50)	0.02 (0.01)	0.02 (0.01)	-0.09
Mother educ: college or more	0.40 (0.49)	0.00 (0.01)	-0.01 (0.01)	0.09
Mother: In labor force (1991)	0.95 (0.22)	0.01 (0.00)	0.00 (0.00)	-0.06
Hotelling's T-squared	129.66			

\*  $p < 0.10$ , \*\*  $p < 0.05$ , \*\*\*  $p < 0.01$

Column 2 and 3 test for differences respectively within school and municipality. Sample restricted to Swedish born students living in the same municipality in grade 3 and 9. Controls: none. Prices adjusted to 2014: 1 SEK = 0.11 €. Standard errors clustered at the school level.

Table 2.A.12:  
Differences in student background by treatment assignment:  
Between-school specification  
1982 cohort - All students

Variable	Control Mean	No controls Difference	Controls Difference	Cohen's d
Female	0.50 (0.50)	-0.00 (0.01)	-0.00 (0.01)	0.01
Foreign born	0.16 (0.37)	-0.04* (0.02)	-0.04* (0.02)	0.11
Swedish born, foreign parent	0.11 (0.32)	-0.01 (0.01)	-0.01 (0.01)	0.03
Divorced parent	0.15 (0.35)	-0.02* (0.01)	-0.02* (0.01)	0.05
Father's Disp. income (1000 kr, age 6 to 9)	207.28 (196.96)	-2.38 (7.69)	-2.27 (7.73)	0.01
Mother's Disp. income (1000 kr, age 6 to 9)	163.85 (63.60)	-4.73** (2.25)	-4.69** (2.32)	0.07
Father with low SES	0.37 (0.48)	0.01 (0.02)	0.00 (0.02)	-0.01
Father with medium SES	0.39 (0.49)	0.00 (0.01)	0.00 (0.01)	-0.01
Father with high SES	0.24 (0.43)	-0.01 (0.02)	-0.01 (0.02)	0.02
Father educ: high school	0.41 (0.49)	0.00 (0.01)	0.00 (0.01)	-0.00
Father educ: college or more	0.34 (0.48)	0.00 (0.03)	0.00 (0.03)	-0.01
Father: Time in labor force	0.90 (0.30)	0.02** (0.01)	0.02** (0.01)	-0.09
Mother with low SES	0.37 (0.48)	0.02 (0.03)	0.02 (0.03)	-0.05
Mother with medium SES	0.48 (0.50)	-0.01 (0.02)	-0.01 (0.02)	0.01
Mother with high SES	0.14 (0.35)	-0.02 (0.02)	-0.02 (0.02)	0.05
Mother educ: high school	0.41 (0.49)	0.02 (0.01)	0.02 (0.01)	-0.05
Mother educ: college or more	0.36 (0.48)	-0.01 (0.03)	-0.01 (0.03)	0.01
Mother: Time in labor force	0.89 (0.32)	0.01 (0.01)	0.01 (0.01)	-0.05
Hotelling's T-squared	40.70			

\*  $p < 0.10$ , \*\*  $p < 0.05$ , \*\*\*  $p < 0.01$

Between-school specification. No sample restriction. The specification with controls includes a control for school size. Prices adjusted to 2014: 1 SEK = 0.11 €. Standard errors clustered at the school level.



Table 2.A.13:  
Differences in student background by treatment assignment:  
Between-school specification  
1982 cohort - Restricted sample

Variable	Control Mean	No controls Difference	Controls Difference	Cohen's d
Female	0.50 (0.50)	-0.00 (0.01)	-0.00 (0.01)	0.01
Foreign born	0.16 (0.37)	-0.02 (0.02)	-0.02 (0.02)	0.11
Swedish born, foreign parent	0.11 (0.32)	-0.00 (0.01)	-0.00 (0.01)	0.03
Divorced parent	0.15 (0.36)	-0.01 (0.01)	-0.01 (0.01)	0.05
Father's Disp. income (1000 kr, age 6 to 9)	207.56 (202.76)	-1.27 (8.69)	-1.10 (8.69)	0.01
Mother's Disp. income (1000 kr, age 6 to 9)	164.19 (64.39)	-2.73 (2.59)	-2.69 (2.64)	0.07
Father with low SES	0.37 (0.48)	-0.00 (0.03)	-0.00 (0.03)	-0.01
Father with medium SES	0.38 (0.49)	0.01 (0.01)	0.01 (0.01)	-0.01
Father with high SES	0.24 (0.43)	-0.00 (0.03)	-0.00 (0.03)	0.02
Father educ: high school	0.41 (0.49)	0.00 (0.01)	0.00 (0.01)	-0.00
Father educ: college or more	0.35 (0.48)	0.01 (0.03)	0.01 (0.03)	-0.01
Father: Time in labor force	0.90 (0.30)	0.02 (0.01)	0.02 (0.01)	-0.09
Mother with low SES	0.37 (0.48)	0.01 (0.03)	0.01 (0.03)	-0.05
Mother with medium SES	0.48 (0.50)	0.00 (0.02)	0.00 (0.02)	0.01
Mother with high SES	0.14 (0.35)	-0.01 (0.02)	-0.01 (0.02)	0.05
Mother educ: high school	0.41 (0.49)	0.01 (0.02)	0.01 (0.02)	-0.05
Mother educ: college or more	0.36 (0.48)	0.00 (0.03)	0.00 (0.03)	0.01
Mother: Time in labor force	0.88 (0.32)	0.00 (0.01)	0.01 (0.01)	-0.05
Hotelling's T-squared	13.34			

\*  $p < 0.10$ , \*\*  $p < 0.05$ , \*\*\*  $p < 0.01$

Between-school specification. Sample restricted to municipalities where less than 85% of the students are in sampled schools. The specification with controls includes a control for school size. Prices adjusted to 2014: 1 SEK = 0.11 €. Standard errors clustered at the school level.

## 2.B Results

### 2.B.1 1972 Cohort

Table 2.B.1: Effect of surveys: Within-municipality specification  
1972 cohort - Swedish born

	All		Low Educ		High Educ	
	No control	Control	No control	Control	No control	Control
Short-term Index	0.02 (0.02)	0.02 (0.02)	0.01 (0.02)	0.01 (0.02)	0.13** (0.05)	0.12** (0.05)
Advanced Math in grade 9	0.02* (0.01)	0.02** (0.01)	0.02 (0.01)	0.01* (0.01)	0.07** (0.03)	0.06** (0.03)
Advanced Eng in grade 9	0.00 (0.01)	0.00 (0.01)	-0.00 (0.01)	-0.00 (0.01)	0.05 (0.04)	0.04 (0.04)
GPA grade 9	0.02 (0.01)	0.01 (0.02)	0.01 (0.01)	0.00 (0.02)	0.11*** (0.03)	0.09*** (0.03)
Long-term Index	0.02 (0.02)	0.01 (0.02)	0.02 (0.02)	0.01 (0.02)	0.06* (0.03)	0.05 (0.03)
Short high school	-0.01 (0.01)	-0.01 (0.01)	-0.00 (0.01)	-0.00 (0.01)	-0.04 (0.02)	-0.03 (0.02)
High school	0.01 (0.01)	0.00 (0.01)	0.00 (0.01)	0.00 (0.01)	0.01 (0.02)	0.01 (0.02)
Short college	0.01** (0.01)	0.01** (0.01)	0.01* (0.01)	0.01* (0.01)	0.04* (0.02)	0.04 (0.02)
College	-0.00 (0.01)	-0.00 (0.01)	-0.01 (0.01)	-0.01 (0.01)	0.01 (0.02)	0.01 (0.02)
Gross income (27-31 yrs old)	2.29 (26.48)	-7.05 (22.74)	4.46 (30.21)	-7.15 (25.34)	-2.34 (60.79)	-6.22 (55.71)
Days/year unemp. (27-31 yrs old)	-1.18 (0.92)	-1.06 (0.96)	-1.06 (1.05)	-0.93 (1.09)	-2.23 (1.90)	-1.94 (1.81)

\*  $p < 0.10$ , \*\*  $p < 0.05$ , \*\*\*  $p < 0.01$

The table presents coefficient and standard error from the regression of each row outcome on a dummy for “being surveyed”, including municipality fixed effects. The main outcomes are standardized indexes for short-term educational attainment and long-term educational attainment and labor market performance. Effects on each component of the index are presented below the index row. GPA is on a 0-5 scale. Income is expressed in 100 SEK, and prices are adjusted to 2014: 1 SEK = 0.11 €. The sample is restricted to Swedish born students with Swedish parents. Standard errors clustered at the school level.

Table 2.B.2: Effect of surveys: Within-school specification  
1972 cohort - Swedish born

	All		Low Educ		High Educ	
	No control	Control	No control	Control	No control	Control
Short-term Index	0.03 (0.02)	0.04* (0.02)	0.01 (0.02)	0.02 (0.02)	0.15*** (0.05)	0.13** (0.05)
Advanced Math in grade 9	0.02** (0.01)	0.03*** (0.01)	0.02 (0.01)	0.02* (0.01)	0.08*** (0.03)	0.07** (0.03)
Advanced Eng in grade 9	0.01 (0.01)	0.01 (0.01)	0.00 (0.01)	0.01 (0.01)	0.08** (0.03)	0.07** (0.03)
GPA grade 9	0.02 (0.02)	0.02 (0.01)	0.01 (0.02)	0.01 (0.02)	0.11** (0.05)	0.09** (0.04)
Long-term Index	0.01 (0.02)	0.02 (0.02)	0.01 (0.02)	0.01 (0.02)	0.04 (0.05)	0.03 (0.05)
Short high school	-0.01 (0.01)	-0.01 (0.01)	-0.00 (0.01)	-0.00 (0.01)	-0.07** (0.03)	-0.07** (0.03)
High school	0.00 (0.01)	0.00 (0.01)	0.00 (0.01)	-0.00 (0.01)	0.03 (0.02)	0.03 (0.02)
Short college	0.02* (0.01)	0.02** (0.01)	0.01 (0.01)	0.01 (0.01)	0.04** (0.02)	0.04** (0.02)
College	-0.01 (0.01)	-0.00 (0.01)	-0.01 (0.01)	-0.00 (0.01)	0.01 (0.02)	0.01 (0.02)
Gross income (27-31 yrs old)	0.42 (28.64)	-2.31 (27.53)	0.67 (32.17)	-6.72 (31.52)	27.07 (72.47)	19.94 (66.56)
Days/year unemp. (27-31 yrs old)	-1.99** (0.83)	-1.86** (0.82)	-1.76** (0.84)	-1.60* (0.84)	-4.00 (2.45)	-3.63 (2.46)

\*  $p < 0.10$ , \*\*  $p < 0.05$ , \*\*\*  $p < 0.01$ 

The table presents coefficient and standard error from the regression of each row outcome on a dummy for “being surveyed”, including school fixed effects. The main outcomes are standardized indexes for short-term educational attainment and long-term educational attainment and labor market performance. Effects on each component of the index are presented below the index row. GPA is on a 0-5 scale. Income is expressed in 100 SEK, and prices are adjusted to 2014: 1 SEK = 0.11 €. The sample is restricted to Swedish born students with Swedish parents. Standard errors clustered at the school level.

Table 2.B.3: Effect of assignment to sampled school:  
1972 cohort - Restricted sample

	All		Low Educ		High Educ	
	No control	Control	No control	Control	No control	Control
Short-term Index	-0.01 (0.05)	-0.03 (0.03)	-0.02 (0.05)	-0.04 (0.03)	0.02 (0.07)	0.03 (0.07)
Advanced Math in grade 9	-0.01 (0.03)	-0.02 (0.02)	-0.01 (0.03)	-0.02 (0.02)	0.00 (0.04)	0.01 (0.04)
Advanced Eng in grade 9	0.00 (0.03)	-0.01 (0.02)	-0.01 (0.03)	-0.01 (0.02)	0.03 (0.04)	0.03 (0.04)
GPA grade 9	-0.01 (0.05)	-0.03 (0.02)	-0.02 (0.04)	-0.03 (0.02)	-0.01 (0.07)	-0.00 (0.06)
Long-term Index	-0.01 (0.05)	-0.02 (0.03)	-0.02 (0.04)	-0.03 (0.03)	-0.03 (0.06)	-0.02 (0.05)
Short high school	0.00 (0.02)	0.01 (0.01)	0.01 (0.02)	0.02 (0.02)	-0.03 (0.04)	-0.03 (0.04)
High school	0.00 (0.01)	0.01 (0.01)	-0.00 (0.01)	0.00 (0.01)	0.03 (0.03)	0.03 (0.03)
Short college	-0.00 (0.01)	-0.01 (0.01)	-0.01 (0.01)	-0.01 (0.01)	0.02 (0.02)	0.02 (0.02)
College	0.01 (0.03)	-0.00 (0.01)	0.01 (0.03)	-0.00 (0.01)	-0.01 (0.02)	-0.01 (0.02)
Gross income (27-31 yrs old)	-2.53 (60.65)	-20.60 (44.93)	2.86 (54.73)	-15.16 (42.34)	-88.46 (110.65)	-76.23 (93.87)
Days/year unemp. (27-31 yrs old)	1.84 (1.54)	1.86 (1.37)	1.99 (1.53)	1.98 (1.37)	1.60 (3.02)	1.48 (2.94)

\*  $p < 0.10$ , \*\*  $p < 0.05$ , \*\*\*  $p < 0.01$

The table presents coefficient and standard error from the regression of each row outcome on a dummy for “studying in a sampled school”. The main outcomes are standardized indexes for short-term educational choices and attainment, and long-term educational attainment and labor market performance. Effects on each component of the index are presented below the index row. All specifications control for school size. Specifications in the “Control” column include controls for family background and parental education. GPA is on a 0-5 scale. Income is expressed in 100 SEK, and prices are adjusted to 2014: 1 SEK = 0.11 €. The sample is restricted to municipalities where less than 85% of the students have been sampled. Standard errors clustered at the school level.

Table 2.B.4: Effect of being surveyed in school:  
1972 cohort - Restricted sample

	All		Low Educ		High Educ	
	No control	Control	No control	Control	No control	Control
Short-term Index	-0.00 (0.18)	-0.11 (0.09)	-0.05 (0.16)	-0.13 (0.09)	0.03 (0.24)	0.05 (0.22)
Advanced Math in grade 9	-0.01 (0.10)	-0.07 (0.07)	-0.04 (0.09)	-0.08 (0.07)	0.02 (0.13)	0.01 (0.13)
Advanced Eng in grade 9	0.02 (0.09)	-0.02 (0.06)	-0.01 (0.08)	-0.03 (0.06)	0.05 (0.13)	0.05 (0.13)
GPA grade 9	-0.06 (0.19)	-0.13 (0.11)	-0.10 (0.16)	-0.15 (0.10)	-0.06 (0.29)	-0.02 (0.27)
Long-term Index	0.07 (0.15)	-0.02 (0.09)	0.02 (0.14)	-0.05 (0.08)	0.15 (0.23)	0.14 (0.21)
Short high school	-0.03 (0.08)	0.00 (0.05)	0.00 (0.08)	0.03 (0.05)	-0.14 (0.13)	-0.15 (0.13)
High school	-0.01 (0.04)	0.00 (0.05)	-0.01 (0.05)	-0.00 (0.05)	0.03 (0.11)	0.04 (0.11)
Short college	0.01 (0.04)	-0.01 (0.04)	-0.01 (0.04)	-0.02 (0.04)	0.08 (0.08)	0.07 (0.08)
College	0.04 (0.09)	-0.01 (0.04)	0.03 (0.09)	-0.01 (0.05)	-0.01 (0.07)	-0.00 (0.07)
Gross income (27-31 yrs old)	198.48 (213.12)	60.16 (167.95)	159.13 (181.62)	39.37 (144.35)	208.66 (521.93)	128.85 (465.93)
Days/year unemp. (27-31 yrs old)	1.28 (5.14)	2.46 (4.57)	3.15 (5.20)	4.06 (4.72)	-7.99 (9.23)	-7.45 (9.10)

\*  $p < 0.10$ , \*\*  $p < 0.05$ , \*\*\*  $p < 0.01$ 

The table presents coefficient and standard error from the regression of each row outcome on a variable measuring the fraction of the school-cohort surveyed. The main outcomes are standardized indexes for short-term educational choices and attainment, and long-term educational attainment and labor market performance. Effects on each component of the index are presented below the index row. All specifications control for school size. Specifications in the “Control” column include controls for family background and parental education. GPA is on a 0-5 scale. Income is expressed in 100 SEK, and prices are adjusted to 2014: 1 SEK = 0.11 €. The sample is restricted to municipalities where less than 85% of the students have been sampled. Standard errors clustered at the school level.

## 2.B.2 1977 Cohort

Table 2.B.5: Effect of surveys: Within-municipality specification  
1972 cohort - Swedish born

	All		Low Educ		High Educ	
	No control	Control	No control	Control	No control	Control
Short-term Index	-0.02 (0.03)	-0.03 (0.02)	-0.02 (0.03)	-0.03 (0.02)	-0.05 (0.05)	-0.06 (0.05)
Advanced Math in grade 9	-0.02 (0.02)	-0.02 (0.02)	-0.02 (0.02)	-0.02 (0.01)	-0.01 (0.04)	-0.01 (0.04)
Advanced Eng in grade 9	-0.00 (0.01)	-0.01 (0.01)	-0.00 (0.01)	-0.01 (0.01)	-0.03 (0.03)	-0.04 (0.03)
GPA grade 9	-0.01 (0.02)	-0.02 (0.02)	-0.00 (0.02)	-0.01 (0.02)	-0.04 (0.05)	-0.04 (0.05)
Long-term Index	-0.00 (0.01)	-0.00 (0.01)	0.00 (0.02)	0.00 (0.02)	-0.06 (0.04)	-0.06 (0.04)
Short high school	-0.00 (0.00)	-0.00 (0.00)	-0.00 (0.00)	-0.00 (0.00)	-0.00 (0.02)	-0.00 (0.02)
High school	0.01 (0.01)	0.02 (0.01)	0.02 (0.01)	0.02* (0.01)	-0.03 (0.04)	-0.03 (0.04)
Short college	-0.01* (0.01)	-0.01* (0.01)	-0.01 (0.01)	-0.01 (0.01)	-0.02 (0.02)	-0.02 (0.02)
College	-0.00 (0.01)	-0.01 (0.01)	-0.01 (0.02)	-0.01 (0.01)	0.00 (0.02)	-0.00 (0.02)
Gross income (27-31 yrs old)	-34.17 (31.35)	-22.79 (27.44)	-29.57 (33.35)	-18.37 (30.40)	-91.77 (63.23)	-83.40 (59.89)
Days/year unemp. (27-31 yrs old)	-1.11* (0.60)	-1.01 (0.62)	-1.20* (0.61)	-1.09* (0.62)	-0.24 (2.54)	-0.33 (2.64)

\*  $p < 0.10$ , \*\*  $p < 0.05$ , \*\*\*  $p < 0.01$ 

The table presents coefficient and standard error from the regression of each row outcome on a dummy for “being surveyed”, including municipality fixed effects. The main outcomes are standardized indexes for short-term educational attainment and long-term educational attainment and labor market performance. Effects on each component of the index are presented below the index row. GPA is on a 0-5 scale. Income is expressed in 100 SEK, and prices are adjusted to 2014: 1 SEK = 0.11 €. The sample is restricted to Swedish born students with Swedish parents. Standard errors clustered at the school level.

Table 2.B.6: Effect of surveys: Within-school specification  
1977 cohort - Swedish born

	All		Low Educ		High Educ	
	No control	Control	No control	Control	No control	Control
Short-term Index	-0.04 (0.03)	-0.04 (0.02)	-0.03 (0.03)	-0.03 (0.02)	-0.06 (0.08)	-0.07 (0.08)
Advanced Math in grade 9	-0.02 (0.02)	-0.01 (0.01)	-0.02 (0.02)	-0.01 (0.01)	0.00 (0.05)	-0.00 (0.05)
Advanced Eng in grade 9	-0.01 (0.01)	-0.01 (0.01)	-0.01 (0.01)	-0.01 (0.01)	-0.02 (0.04)	-0.02 (0.04)
GPA grade 9	-0.01 (0.02)	-0.01 (0.02)	-0.01 (0.02)	-0.01 (0.02)	-0.04 (0.06)	-0.05 (0.06)
Long-term Index	-0.01 (0.02)	-0.00 (0.02)	-0.01 (0.02)	-0.00 (0.02)	-0.04 (0.06)	-0.05 (0.06)
Short high school	-0.01 (0.00)	-0.01 (0.00)	-0.01* (0.00)	-0.01* (0.00)	0.00 (0.02)	0.00 (0.02)
High school	0.03** (0.01)	0.02** (0.01)	0.03** (0.01)	0.03** (0.01)	-0.05 (0.04)	-0.05 (0.04)
Short college	-0.01 (0.01)	-0.01 (0.01)	-0.01 (0.01)	-0.01 (0.01)	-0.01 (0.02)	-0.01 (0.02)
College	-0.01 (0.01)	-0.01 (0.01)	-0.01 (0.01)	-0.01 (0.01)	0.01 (0.03)	0.00 (0.03)
Gross income (27-31 yrs old)	-35.71 (33.68)	-13.92 (30.54)	-35.13 (36.57)	-11.43 (33.54)	-94.15 (87.40)	-98.42 (85.60)
Days/year unemp. (27-31 yrs old)	-1.24 (0.75)	-1.27* (0.76)	-1.23 (0.80)	-1.29 (0.80)	-2.60 (2.70)	-2.30 (2.71)

\*  $p < 0.10$ , \*\*  $p < 0.05$ , \*\*\*  $p < 0.01$ 

The table presents coefficient and standard error from the regression of each row outcome on a dummy for “being surveyed”, including school fixed effects. The main outcomes are standardized indexes for short-term educational attainment and long-term educational attainment and labor market performance. Effects on each component of the index are presented below the index row. GPA is on a 0-5 scale. Income is expressed in 100 SEK, and prices are adjusted to 2014: 1 SEK = 0.11 €. The sample is restricted to Swedish born students with Swedish parents. Standard errors clustered at the school level.

Table 2.B.7: Effect of assignment to sampled school:  
1977 cohort - Restricted sample

	All		Low Educ		High Educ	
	No control	Control	No control	Control	No control	Control
Short-term Index	-0.06 (0.04)	-0.01 (0.03)	-0.06 (0.04)	-0.02 (0.03)	0.04 (0.06)	0.05 (0.06)
Advanced Math in grade 9	-0.04 (0.02)	-0.02 (0.02)	-0.04* (0.02)	-0.03 (0.02)	0.05 (0.04)	0.04 (0.04)
Advanced Eng in grade 9	-0.03 (0.02)	-0.00 (0.02)	-0.02 (0.02)	-0.01 (0.02)	-0.01 (0.04)	0.00 (0.04)
GPA grade 9	-0.01 (0.03)	0.03 (0.02)	-0.01 (0.03)	0.02 (0.02)	0.06 (0.04)	0.06 (0.04)
Long-term Index	-0.01 (0.03)	0.01 (0.02)	-0.01 (0.03)	0.01 (0.02)	0.06 (0.05)	0.04 (0.05)
Short high school	-0.02** (0.01)	-0.02*** (0.01)	-0.01* (0.01)	-0.01** (0.01)	-0.07*** (0.01)	-0.07*** (0.01)
High school	0.02 (0.02)	0.01 (0.01)	0.02 (0.02)	0.01 (0.01)	0.01 (0.03)	0.01 (0.03)
Short college	-0.01 (0.01)	-0.01 (0.01)	-0.01 (0.01)	-0.01 (0.01)	0.02 (0.01)	0.02 (0.01)
College	0.01 (0.02)	0.03** (0.01)	0.01 (0.02)	0.03** (0.01)	0.01 (0.02)	0.01 (0.02)
Gross income (27-31 yrs old)	-50.70 (50.17)	-35.07 (39.74)	-55.79 (50.71)	-34.98 (40.19)	28.90 (92.08)	-38.42 (82.18)
Days/year unemp. (27-31 yrs old)	0.42 (1.02)	0.13 (0.91)	0.88 (0.99)	0.66 (0.90)	-4.00 (2.74)	-3.44 (2.56)

\*  $p < 0.10$ , \*\*  $p < 0.05$ , \*\*\*  $p < 0.01$ 

The table presents coefficient and standard error from the regression of each row outcome on a dummy for “studying in a sampled school”. The main outcomes are standardized indexes for short-term educational choices and attainment, and long-term educational attainment and labor market performance. Effects on each component of the index are presented below the index row. All specifications control for school size. Specifications in the “Control” column include controls for family background and parental education. GPA is on a 0-5 scale. Income is expressed in 100 SEK, and prices are adjusted to 2014: 1 SEK = 0.11 €. The sample is restricted to municipalities where less than 85% of the students have been sampled. Standard errors clustered at the school level.



Table 2.B.8: Effect of being surveyed in school:  
1977 cohort - Restricted sample

	All		Low Educ		High Educ	
	No control	Control	No control	Control	No control	Control
Short-term Index	-0.26*** (0.08)	-0.13** (0.07)	-0.27*** (0.08)	-0.14** (0.07)	-0.05 (0.15)	-0.05 (0.15)
Advanced Math in grade 9	-0.13** (0.05)	-0.08 (0.05)	-0.14*** (0.05)	-0.09* (0.05)	0.04 (0.09)	0.03 (0.09)
Advanced Eng in grade 9	-0.13*** (0.05)	-0.07 (0.04)	-0.13*** (0.05)	-0.07* (0.04)	-0.06 (0.09)	-0.04 (0.09)
GPA grade 9	-0.11 (0.07)	-0.03 (0.06)	-0.11 (0.07)	-0.02 (0.06)	-0.00 (0.09)	-0.07 (0.10)
Long-term Index	0.01 (0.06)	0.06 (0.05)	-0.01 (0.06)	0.04 (0.05)	0.26** (0.12)	0.16 (0.11)
Short high school	-0.04*** (0.01)	-0.05*** (0.01)	-0.03** (0.01)	-0.04*** (0.01)	-0.13*** (0.03)	-0.13*** (0.03)
High school	0.12*** (0.04)	0.06** (0.03)	0.11*** (0.04)	0.05* (0.03)	0.14** (0.07)	0.10 (0.07)
Short college	-0.05*** (0.02)	-0.04*** (0.01)	-0.06*** (0.02)	-0.05*** (0.02)	0.01 (0.03)	0.02 (0.03)
College	0.01 (0.04)	0.08** (0.03)	0.02 (0.04)	0.09** (0.04)	0.03 (0.04)	0.02 (0.04)
Gross income (27-31 yrs old)	-51.22 (99.58)	-20.09 (85.35)	-84.23 (98.81)	-42.78 (85.69)	282.80 (175.47)	140.37 (141.93)
Days/year unemp. (27-31 yrs old)	1.50 (2.35)	1.42 (2.26)	2.93 (2.16)	2.65 (2.08)	-11.64 (7.19)	-6.74 (7.01)

\*  $p < 0.10$ , \*\*  $p < 0.05$ , \*\*\*  $p < 0.01$ 

The table presents coefficient and standard error from the regression of each row outcome on a variable measuring the fraction of the school-cohort surveyed. The main outcomes are standardized indexes for short-term educational choices and attainment, and long-term educational attainment and labor market performance. Effects on each component of the index are presented below the index row. All specifications control for school size. Specifications in the “Control” column include controls for family background and parental education. GPA is on a 0-5 scale. Income is expressed in 100 SEK, and prices are adjusted to 2014: 1 SEK = 0.11 €. The sample is restricted to municipalities where less than 85% of the students have been sampled. Standard errors clustered at the school level.

## 2.B.3 1982 Cohort

Table 2.B.9: Effect of surveys: Within-municipality specification  
1982 cohort - Swedish born, same municipality in grades 3 and 9

	All		Low Educ		High Educ	
	No control	Control	No control	Control	No control	Control
Short-term Index	0.01 (0.02)	0.02 (0.01)	0.01 (0.02)	0.02 (0.01)	-0.04 (0.04)	-0.03 (0.04)
Swedish S.Y. 9 Grade	0.00 (0.01)	0.00 (0.01)	-0.00 (0.01)	0.00 (0.01)	-0.04 (0.04)	-0.03 (0.04)
English S.Y. 9 Grade	0.01 (0.01)	0.01 (0.01)	0.01 (0.01)	0.01 (0.01)	-0.02 (0.04)	-0.01 (0.05)
Swedish S.Y. 9 Grade	0.02 (0.01)	0.02 (0.01)	0.02 (0.01)	0.03* (0.01)	-0.02 (0.04)	-0.02 (0.04)
Long-term Index	0.01 (0.01)	0.01 (0.01)	0.01 (0.01)	0.01 (0.01)	0.00 (0.04)	0.01 (0.04)
Short high school	0.00 (0.00)	0.00 (0.00)	-0.00 (0.00)	-0.00 (0.00)	0.04 (0.03)	0.04 (0.03)
High school	-0.00 (0.01)	-0.00 (0.01)	0.00 (0.01)	0.00 (0.01)	-0.04 (0.05)	-0.04 (0.05)
Short college	-0.00 (0.01)	-0.00 (0.01)	-0.00 (0.01)	-0.00 (0.01)	0.01 (0.03)	0.01 (0.03)
College	0.00 (0.01)	0.00 (0.01)	0.00 (0.01)	0.00 (0.01)	-0.02 (0.02)	-0.02 (0.02)
Gross income (27-31 yrs old)	17.83 (18.81)	14.03 (18.42)	14.96 (19.27)	12.69 (19.02)	42.32 (75.63)	59.70 (68.03)
Days/year unemp. (27-31 yrs old)	-0.61 (0.52)	-0.55 (0.52)	-0.44 (0.57)	-0.42 (0.57)	-1.82 (1.78)	-2.02 (1.76)

\*  $p < 0.10$ , \*\*  $p < 0.05$ , \*\*\*  $p < 0.01$ 

The table presents coefficient and standard error from the regression of each row outcome on a dummy for “being surveyed”, including municipality fixed effects. The main outcomes are standardized indexes for short-term educational attainment and long-term educational attainment and labor market performance. Effects on each component of the index are presented below the index row. Grades are on a 0-4 scale. Income is expressed in 100 SEK, and prices are adjusted to 2014: 1 SEK = 0.11 €. The sample is restricted to Swedish born students with Swedish parents living in the same municipality in grades 3 and 9. Standard errors clustered at the school level.

Table 2.B.10: Effect of surveys: Within-school specification  
1982 cohort - Swedish born, same municipality in grades 3 and 9

	All		Low Educ		High Educ	
	No control	Control	No control	Control	No control	Control
Short-term Index	0.04* (0.02)	0.03 (0.02)	0.03 (0.02)	0.03 (0.02)	-0.00 (0.10)	0.01 (0.10)
Swedish S.Y. 9 Grade	0.01 (0.02)	-0.00 (0.01)	0.00 (0.02)	-0.00 (0.01)	-0.00 (0.08)	0.01 (0.07)
English S.Y. 9 Grade	0.02 (0.02)	0.01 (0.02)	0.02 (0.02)	0.01 (0.02)	0.03 (0.08)	0.03 (0.08)
Swedish S.Y. 9 Grade	0.04** (0.02)	0.03 (0.02)	0.03* (0.02)	0.03 (0.02)	0.02 (0.07)	0.03 (0.07)
Long-term Index	0.01 (0.02)	0.00 (0.02)	0.01 (0.02)	0.01 (0.02)	-0.00 (0.07)	-0.01 (0.07)
Short high school	-0.00 (0.00)	-0.00 (0.00)	-0.01 (0.00)	-0.01 (0.00)	0.03 (0.03)	0.03 (0.03)
High school	-0.01 (0.01)	-0.01 (0.01)	-0.00 (0.01)	-0.00 (0.01)	-0.09* (0.05)	-0.10* (0.05)
Short college	-0.00 (0.01)	-0.01 (0.01)	-0.00 (0.01)	-0.00 (0.01)	0.03 (0.04)	0.03 (0.04)
College	0.02* (0.01)	0.01 (0.01)	0.02 (0.01)	0.01 (0.01)	0.02 (0.03)	0.02 (0.03)
Gross income (27-31 yrs old)	52.09* (31.05)	42.56 (29.59)	48.07 (32.33)	39.63 (30.91)	88.75 (113.20)	65.46 (104.37)
Days/year unemp. (27-31 yrs old)	-1.29* (0.66)	-1.10* (0.66)	-1.28* (0.69)	-1.16* (0.69)	-1.30 (3.03)	-1.28 (2.98)

\*  $p < 0.10$ , \*\*  $p < 0.05$ , \*\*\*  $p < 0.01$

The table presents coefficient and standard error from the regression of each row outcome on a dummy for “being surveyed”, including school fixed effects. The main outcomes are standardized indexes for short-term educational attainment and long-term educational attainment and labor market performance. Effects on each component of the index are presented below the index row. Grades are on a 0-4 scale. Income is expressed in 100 SEK, and prices are adjusted to 2014: 1 SEK = 0.11 €. The sample is restricted to Swedish born students with Swedish parents living in the same municipality in grades 3 and 9. Standard errors clustered at the school level.

Table 2.B.11: Effect of assignment to sampled school:  
1982 cohort - Restricted sample

	All		Low Educ		High Educ	
	No control	Control	No control	Control	No control	Control
Short-term Index	0.04 (0.05)	0.04 (0.04)	0.04 (0.05)	0.04 (0.04)	-0.01 (0.05)	-0.01 (0.05)
Swedish S.Y. 9 Grade	0.01 (0.04)	0.01 (0.03)	0.02 (0.04)	0.02 (0.03)	-0.03 (0.05)	-0.03 (0.05)
English S.Y. 9 Grade	0.06 (0.04)	0.05 (0.03)	0.05 (0.04)	0.05 (0.04)	0.04 (0.05)	0.04 (0.05)
Swedish S.Y. 9 Grade	0.04 (0.04)	0.03 (0.03)	0.03 (0.04)	0.03 (0.03)	0.01 (0.05)	-0.00 (0.04)
Long-term Index	0.02 (0.03)	0.02 (0.02)	0.01 (0.03)	0.01 (0.02)	0.07 (0.05)	0.06 (0.04)
Short high school	0.00 (0.01)	0.00 (0.00)	0.01 (0.01)	0.01 (0.00)	-0.02 (0.02)	-0.02 (0.02)
High school	0.00 (0.02)	0.00 (0.01)	-0.00 (0.02)	-0.00 (0.01)	0.05 (0.03)	0.04 (0.03)
Short college	0.00 (0.01)	-0.00 (0.01)	-0.00 (0.01)	-0.00 (0.01)	0.03** (0.01)	0.04** (0.02)
College	0.02 (0.02)	0.01 (0.01)	0.02 (0.02)	0.02 (0.01)	-0.00 (0.02)	-0.01 (0.02)
Gross income (27-31 yrs old)	34.23 (48.32)	26.70 (40.22)	31.83 (49.03)	27.42 (43.24)	36.32 (76.19)	21.61 (70.47)
Days/year unemp. (27-31 yrs old)	0.65 (1.15)	0.78 (0.85)	1.16 (1.00)	1.15 (0.82)	-2.30 (2.64)	-1.67 (2.47)

\*  $p < 0.10$ , \*\*  $p < 0.05$ , \*\*\*  $p < 0.01$

The table presents coefficient and standard error from the regression of each row outcome on a dummy for “studying in a sampled school”. The main outcomes are standardized indexes for short-term educational choices and attainment, and long-term educational attainment and labor market performance. Effects on each component of the index are presented below the index row. All specifications control for school size. Specifications in the “Control” column include controls for family background and parental education. Grades are on a 0-4 scale. Income is expressed in 100 SEK, and prices are adjusted to 2014: 1 SEK = 0.11 €. The sample is restricted to municipalities where less than 85% of the students have been sampled. Standard errors clustered at the school level.

Table 2.B.12: Effect of being surveyed in school:  
1982 cohort - Restricted sample

	All		Low Educ		High Educ	
	No control	Control	No control	Control	No control	Control
Short-term Index	-0.09 (0.07)	0.04 (0.06)	-0.07 (0.08)	0.05 (0.06)	-0.07 (0.10)	-0.04 (0.10)
Swedish S.Y. 9 Grade	-0.07 (0.07)	0.03 (0.06)	-0.05 (0.07)	0.05 (0.06)	-0.08 (0.09)	-0.06 (0.09)
English S.Y. 9 Grade	-0.01 (0.07)	0.08 (0.06)	-0.02 (0.08)	0.07 (0.06)	0.14 (0.10)	0.14 (0.10)
Swedish S.Y. 9 Grade	-0.06 (0.06)	0.01 (0.05)	-0.06 (0.06)	0.02 (0.05)	-0.03 (0.09)	-0.06 (0.09)
Long-term Index	-0.04 (0.05)	-0.01 (0.04)	-0.07 (0.05)	-0.03 (0.03)	0.22** (0.10)	0.13 (0.10)
Short high school	0.01 (0.01)	0.00 (0.01)	0.01 (0.01)	0.01 (0.01)	-0.03 (0.03)	-0.02 (0.03)
High school	0.11*** (0.03)	0.07*** (0.02)	0.10*** (0.03)	0.06*** (0.02)	0.13** (0.06)	0.09 (0.06)
Short college	-0.03** (0.02)	-0.03* (0.01)	-0.03* (0.02)	-0.03* (0.02)	-0.02 (0.03)	-0.01 (0.03)
College	-0.02 (0.03)	0.03 (0.02)	-0.02 (0.03)	0.03 (0.02)	0.05 (0.04)	0.05 (0.04)
Gross income (27-31 yrs old)	-81.90 (80.30)	-76.43 (67.52)	-106.97 (82.21)	-85.16 (72.16)	141.69 (168.89)	-27.45 (168.29)
Days/year unemp. (27-31 yrs old)	3.71* (2.04)	3.64** (1.65)	5.50*** (1.88)	5.14*** (1.64)	-9.62* (5.19)	-5.40 (5.04)

\*  $p < 0.10$ , \*\*  $p < 0.05$ , \*\*\*  $p < 0.01$

The table presents coefficient and standard error from the regression of each row outcome on a variable measuring the fraction of the school-cohort surveyed. The main outcomes are standardized indexes for short-term educational choices and attainment, and long-term educational attainment and labor market performance. Effects on each component of the index are presented below the index row. All specifications control for school size. Specifications in the “Control” column include controls for family background and parental education. Grades are on a 0-4 scale. Income is expressed in 100 SEK, and prices are adjusted to 2014: 1 SEK = 0.11 €. The sample is restricted to municipalities where less than 85% of the students have been sampled. Standard errors clustered at the school level.

## References

- Altonji G. John, 1993. "The Demand for and Return to Education when Education Outcomes are Uncertain," *Journal of Labor Economics*, 48–83.
- Anderson L. Michael, 2008. "Multiple Inference and Gender Differences in the Effects of Early Intervention: A Reevaluation of the Abecedarian, Perry Preschool, and Early Training Projects," *Journal of the American statistical Association* 103(484).
- Angrist D. Joshua & Lavy V., 1999. "Using Maimonides' rule to Estimate the Effect of Class Size on Scholastic Achievement," *Quarterly Journal of Economics* 114 (2).
- Arcidiacono Peter, 2004. "Ability Sorting and the Returns to College Major," *Journal of Econometrics* 121(1), 343–375.
- Arcidiacono Peter & Aucejo E. M. & Fang H. & Spenner I. K., 2011. "Does Affirmative Action Lead to Mismatch? A New Test and Evidence," *Quantitative Economics* 2 (3), 303–333.
- Arcidiacono Peter & Hotz J. V. & Kang S., 2012. "Modeling College Major Choices Using Elicited Measures of Expectations and Counterfactuals," *Journal of Econometrics* 166 (1), 3–16.
- Avvisati Francesco & Gurgand M. & Guyon N. & Maurin E., 2014. "Getting Parents Involved: A Field Experiment in Deprived Schools," *The Review of Economic Studies* 81 (1), 57–83.
- Bailar A. Barbara, 1975. "The Effects of Rotation Group Bias on Estimates from Panel Surveys," *Journal of the American Statistical Association* 70 (349), 23–30.
- Barberis Nicholas & Thaler R., 2003. "A Survey of Behavioral Finance," *Handbook of the Economics of Finance* 1, 1053–1128.
- Björklund Anders & Clark A. Melissa & Edin Per-Anders & Fredriksson Peter & Krueger B. Alan , 2005. "The Market Comes to Education in Sweden: an evaluation of Sweden's surprising school reforms," Russell Sage Foundation Publications.

Bound John & Brown C. & Mathiowetz N., 2001. "Measurement Error in Survey Data," *Handbook of econometrics* 5, 3705–3843.

Cunha Flavio & Heckman J., 2007. "The Technology of Skill Formation," *The American Economic Review* 97(2), 31–47.

Cunha Flavio & Heckman J. & Schennach M. S., 2010. "Estimating the Technology of Cognitive and Noncognitive Skill Formation," *Econometrica* 78 (3), 883–931.

DellaVigna Stefano, 2009. "Psychology and Economics: Evidence from the Field," *Journal of Economic Literature* 47(2), 315–372.

Duflo Esther & Zwane P. A. & Zinman J. & Van Dusen E. & Pariente W. & Null C. & Miguel E. & Kremer M. & Karlan S. D. & Hornbeck R. & Giné X., 2011. "Being Surveyed can Change Later Behavior and Related Parameter Estimates," *Proceedings of the National Academy of Sciences* 108(5), 1821–1826.

Emanuelsson Ingemar, 1979. "Utvärdering Genom Uppföljning av Elever ett Nytt Individual-statistikprojekt".

Facchinello Luca, 2016. "The Impact of Early Grading on Academic Choices: Mechanisms and Social Implications," working paper, Stockholm School of Economics.

Facchinello Luca, 2016. "Does Peer Ability Affect Education Choices?," working paper, Stockholm School of Economics.

Fredriksson Peter & Öckert B. & Oosterbeek H., 2013. "Long-term Effects of Class Size," *Quarterly Journal of Economics* 128(1), 249–285.

Fryer G. Roland Jr & Levitt D. S. & List A. J., 2015. "Parental Incentives and Early Childhood Achievement: A Field Experiment in Chicago heights," Technical report, National Bureau of Economic Research.

Goux Dominique & Gurgand M. & Maurin E., 2014. "Adjusting your Dreams? High school Plans and Dropout Behavior," *IZA discussion paper* No. 7948.

Härnqvist Kjell, 1998. "A Longitudinal Program for Studying Education and Career Development".

Heckman James & Moon S. H. & Pinto R. & Savelyev P. & Yavitz A., 2010. "Analyzing Social Experiments as Implemented: A Reexamination of the Evidence from the Highscope Perry Preschool Program," *Quantitative Economics* 1(1), 1–46. 25

Heckman James & Mosso S., 2014. "The Economics of Human Development and Social Mobility," *Annual Review of Economics* 6(1), 689–733.

Heckman James & Pinto R. & Savelyev P., 2013. "Understanding the Mechanisms through which an Influential Early Childhood Program Boosted adult Outcomes," *The American Economic Review* 103(6), 1–35.

Heckman James & Pinto R. & Shaikh M. A. & Yavitz A., 2011. "Inference with Imperfect Randomization: The case of the Perry Preschool Program," Technical report, National Bureau of Economic Research.

Jalava Nina & Joensen J. S. & Pellas E., 2015. "Grades and Rank: Impacts of Non-financial Incentives on Test Performance," *Journal of Economic Behavior & Organization* 115, 161–196.

Jensen Robert, 2010. "The (perceived) Returns to Education and the Demand for Schooling," *Quarterly Journal of Economics* 125(2), 515–548.

Joensen S. Juanna & Nielsen H. S., 2015. "Spillovers in Educational Choice," SSRN 2548702.

Kling R. Jeffrey & Liebman B. J. & Katz F. L., 2007. "Experimental Analysis of Neighborhood Effects," *Econometrica* 75(1), 83–119.

Lee Soohyung & Shaikh M. A., 2014. "Multiple Testing and Heterogeneous Treatment Effects: Re-evaluating the Effect of Progreso on School Enrollment," *Journal of Applied Econometrics* 29(4), 612–626.

Lehmann L. Erich & Romano P. J., 2006. "Testing Statistical Hypotheses," Springer Science & Business Media.

List A. John & Shaikh M. A. & Xu Y., 2016. "Multiple Hypothesis Testing in Experimental Economics," Technical report, National Bureau of Economic Research.



Nguyen Trang, 2008. "Information, Role Models and Perceived Returns to Education: Experimental Evidence from Madagascar," MIT mimeo. 26.

Romano P. Joseph & Wolf M., 2005a. "Exact and Approximate Stepdown Methods for Multiple Hypothesis Testing," *Journal of the American Statistical Association* 100(469), 94–108.

Romano P. Joseph & M. Wolf, 2005b. "Stepwise multiple testing as formalized data snooping," *Econometrica* 73(4), 1237–1282.

Romano P. Joseph & Wolf M., 2010. "Balanced Control of Generalized Error Rates," *The Annals of Statistics*, 598–633.

Simon A. Herbert, 1955. "A Behavioral Model of Rational Choice," *Quarterly Journal of Economics*, 99–118.

Solon Gary, 1986. "Effects of Rotation Group Bias on Estimation of Unemployment," *Journal of Business & Economic Statistics* 4 (1), 105–109.

Westfall H. Peter & Young S. S., 1993. "Resampling-based Multiple Testing: Examples and methods for p-value adjustment," Volume 279. John Wiley & Sons.

Wiswall Matthew & Zafar B., 2015. "Determinants of College Major Choice: Identification using an information experiment," *The Review of Economic Studies* 82(2), 791–824.

Zafar Basit, 2011. "How do College Students Form Expectations?," *Journal of Labor Economics* 29(2), 301–348.



## Chapter 3

# Does Peer Ability Affect Education Choices?<sup>1</sup>

---

<sup>1</sup>A special thanks to my two advisors, Erik Lindqvist and Juanna Joensen, who provided helpful comments throughout the paper.

### 3.1 Introduction

Most of the literature on peer effects in the economics of education focuses on how peer performance affects individual school achievement. The literature, recently reviewed by Sacerdote (2011), finds either positive, zero or non-linear effects.<sup>2</sup>

To make sense of the mixed results recent work has started to focus more explicitly on mechanisms. An emerging literature highlights how better peers can negatively affect student performance and academic choices due to rank concerns. Murphy & Weinhardt (2014) find that students who rank high in their school develop higher confidence, and perform better in class. Tincani (2015) shows that rank concerns and academic competition generate positive peer effects in academic performance among Chilean 8th graders. Elsner & Isphording (2015) find that ranking higher in high school significantly affects perceived intelligence, teacher support and long-run educational outcomes of students.

In this paper I investigate how classmates' ability affects the school choices of Swedish 6th graders. Instead of considering only peer effects in performance or pure rank effects, I look at the overall effect of peer ability on education choices. I then proceed to study underlying mechanisms.

The institutional setup is particularly suitable to answer the research question. Swedish students do not get formal grades during the academic year, and start to be assigned end-of-the-year grades relatively late during compulsory school (typically in grade 7, at age 14). Class interactions might play an important role for the formation of beliefs about own ability in such an environment. At the same time students need to assess their academic ability when they choose whether to take advanced or general courses in grades 7 to 9. Advanced courses are more challenging and prepare for academic high school tracks, the only

---

<sup>2</sup> Hanushek et al. (2003), Ammermueller & Pischke (2009) and Imberman et al. (2012) find positive peer effects on student performance in primary or secondary school. Sund (2009), Lavy et al. (2012a), Burke & Sass (2013), Abdulkadiroglu et al. (2014) and Tincani (2015) find instead non-linear or zero effects.

option to access college.

These choices are the main outcomes in my analysis. They are recorded in my dataset, together with IQ-like ability tests, for a representative sample of Swedish 6th grade classes.

I identify the effect of interest by exploiting variation in average class ability within schools. I show that this variation is orthogonal to different measures of own ability, parental occupation, special education status, gender, and class size. This is evidence of students being allocated to classes independently of their ability, which is credible given school administrators lack early measures of academic performance.

I find that a one standard deviation increase in average class ability reduces by 2 percentage points the probability of taking an advanced math course in grades 7 to 9. Peer ability does not affect English course choices in grades 7 to 9, and whether students choose academic tracks in high school.

Children's survey responses and administrative data allow me to look at the underlying mechanisms. First, I find that peer ability negatively affects students' assessment of own ability. Second, I look at how peer ability affects academic performance. I show that students benefit much more from having high ability peers in English, a subject involving group interaction, than they do in math. Assuming that academic performance positively affects education choices, this could explain why I find a negative effect of peer ability on course choices only in math. Finally peer ability does not seem to affect student's motivation, class interaction and parental support, but positively affects teacher interaction.

In terms of mechanisms, results are consistent with Elsner & Isphording (2015), who find that higher school rank improves perceived intelligence and teacher support. The same paper however also finds a positive effect on long-run academic choices and performance. I find no evidence of an overall increase in student confidence ("motivation" in my analysis), as in Murphy & Weinhardt (2014), or lower academic performance, as predicted by Tincani's model of rank concerns. I cannot however exclude that these two effects might be at play, when keeping ability fixed. All the cited papers indeed look at a pure rank effect. If peer ability affects performance in the opposite direction, as shown in my results, the overall effect on performance and motivation will systematically differ.

The main contribution of the paper is indeed estimating the net effect of

peer ability on academic choices. Ultimately this is the relevant parameter policy makers should consider when designing tracking and streaming policies.

The paper proceeds as follows. In Section 3.2 I describe the data, the education system, and present descriptives. In Section 3.3 I discuss potential mechanisms and identification, and present the main results. Section 3.4 presents additional results related to the mechanisms. Section 3.5 draws conclusions.

## 3.2 Setup

### 3.2.1 Data

I use data from *Evaluation Through Follow-up* (ETF), a longitudinal project surveying every five years representative samples of Swedish students enrolled in compulsory school. I use wave three of the study, which sampled whole classes of 6th graders from 29 municipalities.<sup>3</sup> Most students were born in 1967, and were around 13 at the time of sampling. While classes obviously change over time, all students, apart from those who moved abroad, were followed up.

The most important piece of information in my analysis, the “treatment”, is the cognitive ability of the students. During the spring term of school year 6 most students (89% of the final sample) took a battery of three standardized ability tests: a test of verbal ability, a test of inductive ability, and a test of spatial ability. Students had respectively 15, 27 and 22 seconds to answer each section of the test, assuming they wasted no time at all. The fast pace of the test adds to the quality of the ability measures: Borghans et al. (2008) show that reducing the time available for completing intelligence tests reduces differences in effort between students with different non-cognitive traits. Moreover at the time of the tests students were 13, a point in which IQ should have already stabilized (Cunha & Heckman, 2009).

A unique feature of the data is the detailed survey that most children (92.5% of the final sample) filled in in grade 6. Among the questions asked, children

---

<sup>3</sup>Municipalities were drawn using stratified sampling. Strata are defined by population, fraction of left-wing voters, fraction working in the public sector and fraction of immigrants. The three biggest municipalities in Sweden (Stockholm, Malmö, Gothenburg) are always part of the sample. Further details on the sampling scheme can be found in Emanuelsson (1979).

had to evaluate own ability in different domains, and were asked to rank themselves in the class. The same survey also asked many questions about school motivation, parental support at home, perceived teacher helpfulness, and class environment. These questions allow me to test *how* peer ability affects education choices, and thus to better clarify mechanisms.

Finally the ETF data contains basic demographics taken from Statistics Sweden (gender, parental occupation, socioeconomic status, foreign status) and detailed school level data up to college (course choices, absences, changes of teachers and grades in compulsory school, tracks and grades in high school). This allows me to look both at choices taken during grade 6 (course choices in grades 7) and choices taken some years after (course choices in grades 8 and 9, track in high school). Test scores data allows me to inspect potential spillovers in performance of class ability.

The original sample consists of roughly 9000 Swedish compulsory school students (10% of the targeted population) living in 29 (out of 290) municipalities, the lowest administrative unit in Sweden. I exclude from my sample classes with less than 10 students, which are more likely to cater exclusively to special education students. This reduces the sample by 8%. About 15% of the students misses ability data, so the final sample consists of 7913 students, 373 classes, and 257 schools. My preferred specification uses school fixed effects. This sample, from now on the *restricted sample*, uses only schools where more than one class was sampled. It consists of 4452 students, 204 classes, and 88 schools (50% of the final sample).

### 3.2.2 Education System

Compulsory school in Sweden starts at age seven, and lasts nine years. In this period it was formally divided in three stages, which could also entail students changing schools: early compulsory school (grades 1-3), middle compulsory school (grades 4-6), and late compulsory school (grades 7-9). While the data does not report whether students were in the same class starting from grade 1, the majority of sampled students should have stayed for at least 3 years in the same class, from age 10 to 13.

Whether students rely on peer interaction to assess own ability depends also on the feedback information they receive in class. In Sweden students are not assigned formal grades during the school year. They instead receive homework and test scores (without a pass or fail threshold). End of the year formal grades,

based on standardized national tests, were assigned to all students in grades 7-9. About half of the sample received grades also at the end of school year 6, and, less likely, of grade 3. The decision to assign early grades was taken by municipal school boards.<sup>4</sup>

What is relevant for the analysis is that up to the grade considered students had only a limited amount of information to assess their ability, especially in relation to a national benchmark, or a proper criterion. Classroom peers might thus have played a bigger role in how these students were forming beliefs about own ability.

In the spring of school year 6, children chose whether to take math and English at the advanced or general level in the next school year. Advanced courses provided better preparation for academic tracks in high school. Students were allowed to switch course type over time - the most common switch was from an advanced to a general course. At the end of compulsory school, students could enroll in either academic or vocational high school tracks. Vocational tracks lasted two years, provided professional training, and did not allow direct access to college. Academic high school tracks lasted three or four years, prepared for college, and could be selective.<sup>5</sup>

After academic high school graduation (or after taking one more year of high school after vocational school) students became eligible to apply to college. A student quota, set by the government, limited access to college. Slots were competitively assigned to the students with highest GPA or *SweSAT* (a college entry test similar to the American SAT).<sup>6</sup> College was tuition-free, and a mix of grants and income-contingent loans allowed admitted students to pay for living expenses. Higher education was thus both meritocratic and competitive.

---

<sup>4</sup>Sjögren (2010) provides details on the implementation of the reform that abolished “early grades” - grades in school years 3 and 6. Facchinello (2016) focuses on the cohort born 1967 to evaluate the effects of the reform in the short and long-run.

<sup>5</sup>A high grade 9 GPA and advanced math electives in compulsory school could be used as admission requirement.

<sup>6</sup>Öckert (2002) reports that around 50% of the students were rejected admission to college in the period I study, confirming the selective nature of Swedish higher education.



### 3.2.3 Descriptives

Table 3.2.1 shows ability, demographics and class level information for the final and *restricted sample*. The sample restricted to schools where more than one class was sampled appears to be very similar to the final sample. The only differences are a slightly lower level of ability and marginally larger classes. 40% of the students come from a working class family (coded as low SES), and around 13% receive special education. The fraction of students with foreign parents is quite low, 6%. There is variation also at the class level, which appears to be particularly the case for class average ability, the share of students who are immigrant, low SES and special education (see columns four and eight). This implies that the education system is not uniform, with some segregation at the class level.

Table 3.2.1: Background: whole and restricted sample

	All sample				Final sample			
	Mean	Median	Sd	Class Sd	Mean	Median	Sd	Class Sd
<b>Ability</b>								
Standardized verbal ability	0.00	0.01	1.00	0.36	-0.02	0.01	1.00	0.37
Standardized inductive ability	-0.00	0.08	1.00	0.38	-0.01	-0.04	0.99	0.35
Standardized general ability	0.00	0.06	1.00	0.39	-0.02	0.06	0.99	0.37
<b>Demographics</b>								
Male	0.51	1.00	0.50	0.10	0.52	1.00	0.50	0.12
Low SES	0.40	0.00	0.49	0.17	0.40	0.00	0.49	0.17
Parent not Nordic	0.06	0.00	0.24	0.06	0.06	0.00	0.23	0.06
Special education	0.13	0.00	0.34	0.12	0.13	0.00	0.34	0.12
<b>Class</b>								
Changes of teacher	0.50	0.00	0.81	0.82	0.51	0.00	0.84	0.84
Class size	24.00	25.00	4.46	4.87	24.54	25.00	3.61	3.91

The restricted sample includes only those schools where multiple classes were sampled. Class sd is the standard deviation of the class average. Ability measures are IQ-like scores taken at the end of grade 6. General ability is the average of inductive and verbal ability.

Table 3.2.2 presents descriptives for the main outcomes in the analysis. In terms of school choices, about three quarters of the students select advanced courses in grade 7. However over time many switch to the easier general courses. This pattern is more pronounced for math. 85% of the sample proceeds into high school education. Of the enrolled students, 56% choose an academic track,

which grants eligibility, but not necessarily admission, to college.

Apart from the English and math test scores, all variables in the second part of the table use the answers given by students in the grade 6 questionnaire, filled in by 92.6% of the sample.

Students were asked to assess their skills in the class on a 1-9 scale, with 9 representing the top in the class. I interpret this as a measure of *relative evaluation of skills*. Apparently students tend to over-rank themselves when they evaluate their skills: the distribution would be normal and centered on the mean value if evaluations were unbiased. *Evaluation of skills* is an index derived from yes/no questions asking students whether they are good at math, spelling, reading and at school in general. While the average student answers positively to most questions, there is quite some variation in the answers.

The English and Math scores are the results of the national tests, and determine end-of-the-year grades. Notice the substantial amount of variation at the class level.

*Motivation*, *Class environment* and *Teacher helpfulness* are indexes built from survey questions asking respectively students: to evaluate own ability in different domains (math, spelling, reading, etc), their motivation in school (if they put effort in school tasks, if they give up easily, etc), availability and helpfulness of the teacher, and how they get along with their classmates. While *teacher helpfulness* and especially *class environment* are rated highly by most of the students, there appears to be more variation in student *motivation*. Finally most students appear to receive help at home with homework, though there is some variation at the class level.<sup>7</sup>

---

<sup>7</sup>Appendix 3.A presents the distribution of the discrete variables and provides details on the constructions of the indexes used in the analysis.

Table 3.2.2: Outcomes: whole and restricted sample

	All sample				Final sample			
	Mean	Median	Sd	Class Sd	Mean	Median	Sd	Class Sd
<b>Choices</b>								
Advanced math in grade 7	0.74	1.00	0.44	0.14	0.75	1.00	0.43	0.13
Advanced math in grade 8	0.66	1.00	0.47	0.15	0.67	1.00	0.47	0.14
Advanced math in grade 9	0.55	1.00	0.50	0.16	0.56	1.00	0.50	0.17
Advanced English in grade 7	0.77	1.00	0.42	0.13	0.77	1.00	0.42	0.13
Advanced English in grade 8	0.74	1.00	0.44	0.13	0.74	1.00	0.44	0.13
Advanced English in grade 9	0.69	1.00	0.46	0.14	0.69	1.00	0.46	0.15
Academic track in high school	0.56	1.00	0.50	0.17	0.55	1.00	0.50	0.16
<b>Mechanisms</b>								
Evaluation of skills in class (1-9)	5.75	6.00	1.59	0.42	5.77	6.00	1.58	0.42
Evaluation of skills (1-9)	6.68	7.00	2.43	0.77	6.71	7.00	2.42	0.76
Math std test score, Grade 9	0.01	0.06	1.00	0.52	0.01	0.06	0.97	0.53
English std test score, Grade 8	0.01	0.11	1.00	0.43	0.01	0.11	0.99	0.46
Motivation in school (0-9)	5.66	6.00	2.05	0.62	5.68	6.00	2.04	0.62
Class environment (1-9)	7.91	9.00	1.65	0.48	7.93	9.00	1.62	0.57
Teacher helpfulness (0-9)	7.07	7.50	1.86	0.62	7.03	7.50	1.87	0.67
Help at home with homework	0.80	1.00	0.40	0.12	0.80	1.00	0.40	0.12

The restricted sample includes only those schools where multiple classes were sampled. Class sd is the standard deviation of the class average. *Evaluation of skills in class* is the answer to a question asking where the student is in terms of skills in the class distribution. *Evaluation of skills*, *Motivation*, *Class environment* and *Teacher helpfulness* are indexes built from survey questions asking respectively students: to evaluate own ability in different domains (math, spelling, reading, etc), their motivation in school (if they put effort in school tasks, if they give up easily, etc), availability and helpfulness of the teacher, and how they get along with their classmates.

### 3.3 Empirics

In this section I first discuss the mechanisms through which I expect peer ability to influence individual choices. I then consider identification issues, and finally present the main results.

#### 3.3.1 Mechanisms

The empirical model I would like to estimate is the following:

$$Y_{ics} = \alpha + \beta \text{Ability}_{-i,c} + \gamma \text{Ability}_i + \epsilon_{ics} \quad (3.1)$$

where  $i$  indexes the student,  $c$  the classroom, and  $s$  the school. I am interested in  $\beta$ , the reduced form effect of peer ability on individual choices. While peer ability might affect students in many different ways, the net effect might be more interesting for a policy maker when considering class formation and streaming policies.

Different mechanisms might indeed contribute to the final effect. Among the ones that might affect positively education choices there are performance spillovers and peer imitation. Higher peer ability implies higher peer performance. Many studies find positive spillover effects from being surrounded by better performing peers.<sup>8</sup> Positive peer effects in performance translate into a lower cost of education. In a simple model of education choice this leads to more investment in education.

Higher peer ability also implies more students choosing academic education. If there are imitation or role model effects, students' education choices would be positively affected.

On the negative side students could make inferences about own ability using the class as a benchmark. They could for instance judge their ability in relation to their rank in the class.<sup>9</sup> Average ability students surrounded

---

<sup>8</sup>Studies that find positive peer effects on student performance in primary or secondary school are Hanushek et al. (2003), Ammermueller & Pischke (2009), Imberman et al. (2012). Studies that find a non-linear or zero effect are Sund (2009), Lavy et al. (2012), Burke & Sass (2013), Abdulkadiroglu et al. (2014) and Tincani (2015).

<sup>9</sup>Recent literature shows that children have imperfect knowledge of their ability. See for in-

by high ability peers rank low in the class, might conclude that they are low ability, and thus revise downward their education choices. The opposite would be true for students who interact with low ability classmates. This might be particularly important in the educational setup here considered, as most students up to grade 6 do not have any formal grades. In appendix 3.B I show that student ability percentile rank within the class can differ substantially from the national percentile rank (a more objective measure of ability). Moreover student percentile rank can differ substantially between classes.

Secondly, peer ability might affect student motivation. This is related, but different from the point above: even if students are not “framed” by their peers’ ability, they might feel less enthusiastic about school when their performance is systematically lower than the one of their peers. To what extent this might be true depends on both the education setup (how competitive the education system is) and social norms (to what extent people are affected by social comparisons). Swedish school curricula specifically emphasize cooperation and integration, at the expense of class competition, so this effect might be less important here.

Other mechanisms that might affect education choices in less clear ways are responses from classmates, teachers and parents. A class with higher ability might have less disruptive kids, and thus offer a better learning environment. At the same time, being surrounded by higher ability peers might create frictions between the children and their peers, for instance due to a more competitive environment. In higher ability classes teachers might make classes more challenging, but could also devote more time to low ability students. Finally it is not clear whether parents are substitutes or complements in the education production function, so they might either counteract or amplify any of the effects listed above.

Ultimately which of these mechanisms prevails is an empirical question. In my analysis I will be able to look separately at most of the mechanisms I mentioned.

---

stance Stinebrickner & Stinebrickner (2012, 2014), Zafar (2011), Bobba & Frisanchio (2014), Facchinello (2016). Murphy & Weinhardt (2014), Tincani (2015) and Elsner & Isphording (2015) find that school and class rank matters for performance and education choices.

### 3.3.2 Identification

The effect I aim to estimate is called in the peer effects literature a “contextual effect”: the effect of a pre-determined peer characteristic on an individual outcome. Identification of this effect requires exogenous, or conditionally exogenous variation in peer ability.

If peer ability is randomly assigned at the cohort level, then  $\beta$  in equation 3.1 has a causal interpretation. This is likely not to be the case, as peer ability is likely to correlate with many other factors that also affect education choices (own ability, SES, school quality, etc). However, if school administrators do not observe ability when they form classes, variation in peer ability might still be random within school. This is likely to be the case in Sweden, where there are no formal grades (proxying for ability) in early stages of education. My identification strategy relies on this exogenous variation. My final specification is:

$$Y_{ics} = \alpha + \beta \text{Ability}_{-i,c} + \gamma \text{Ability}_i + \text{School}_s + \Delta X_{ics} + \epsilon_{ics} \quad (3.2)$$

where  $X_{ics}$  contains controls at the individual level.

Two issues specific to the identification of peer effects are “multiplier effects” (because individual  $i$ ’s behavior is affected by peer behavior, peer behavior will also be affected by  $i$ ’s behavior) and correlated shocks (rather than a genuine peer effect, the estimated effect might reflect some specific shock common to the group, in this case the class).

If peer ability affects students’ choices at ages 14 to 16, those choices cannot have an effect on the ability of the students, measured when students were 13. It is still possible that peer ability in past grades affected students’ own characteristics, like own ability. In this case I estimate the effect of “final peer ability” on later education choices, disregarding the effect of peer ability on own ability in earlier stages. This is not a threat to identification, but affects the interpretation of my estimates.

In terms of correlated shocks, the choices of the students and the ability composition of the class are not measured at the same time, so the problem should not be there. However class level characteristics, for instance teachers, might affect both choices and peer ability. The literature (e.g., Heckman et al., 2007) seems to support the idea that ability at age 13 is stable, so I can exclude this type of effects.

Another important point is that ability measures typically contain measurement error. In my specific case I use an average measure of ability, so there should be very little measurement error. As long as peer ability is uncorrelated with own ability, which instead is going to be affected by measurement error, there should be no bias in  $\beta$ .

In Tables 3.3.1 and 3.3.2 I test whether peer ability affects predetermined factors that might affect education choices. Column 1 and 2 report respectively the coefficient  $\beta$  from the following two regressions:

$$Y_{ics} = \alpha + \beta \text{Ability}_{-i,c} + \epsilon_{ics} \quad (3.3)$$

$$Y_{ics} = \alpha + \beta \text{Ability}_{-i,c} + \gamma \text{Ability}_{-i,s} + \text{School}_s + \epsilon_{ics} \quad (3.4)$$

Equation 3.4 adds school fixed effects to Equation 3.3, and a control for average ability of peers in the same school, excluding  $i$ . This test correction has been suggested by Guryan et al. (2009), who show that in tests of peer ability a mechanical negative bias is induced by the fact that higher peer ability in a group implies lower own ability within the group.

Column 1 of Table 3.3.1 confirms the selection problem mentioned above: higher peer ability is positively correlated with higher own ability, kindergarten attainment, SES (which is reflected in parental occupation in Table 3.3.2), and is negatively correlated with having foreign parents and teacher turnover.

Column 2 of Tables 3.3.1 and 3.3.2 shows that within-school variation in peer ability is not correlated with these variables. However I still find that one occupational dummy out of ten, *higher civil servants and senior salaried*, is positively associated with peer ability. This suggests that parents with high social status might still be able to place their children in better classes.<sup>10</sup>

An important caveat is that I have no data on teachers, so I cannot test whether more qualified or experienced teachers are systematically assigned to classes with higher or lower ability. It is however reassuring to see that the test for teacher turnover passes. This suggests that teacher assignment does not change with class ability.

---

<sup>10</sup>In the results section I perform robustness tests to understand whether this potential deviation from random assignment affects the coefficient of interest.

Table 3.3.1:  
Does peer ability predict student background?

Outcome	OLS	School FE
Standardized general ability	0.68*** (0.02)	0.02 (0.03)
Standardized verbal ability	0.56*** (0.03)	0.06 (0.04)
Standardized inductive ability	0.62*** (0.03)	0.00 (0.04)
Male	-0.01 (0.01)	-0.02 (0.02)
Parent not Nordic	-0.01* (0.01)	0.02 (0.01)
Low SES	-0.17*** (0.02)	-0.03 (0.03)
Kindergarten	0.04** (0.02)	0.01 (0.02)
Hours of absence in grade 6	11.61 (56.48)	-53.28 (57.56)
Special education	-0.02 (0.02)	0.02 (0.02)
No ability data	0.05 (0.03)	0.08 (0.05)
Changes of teacher	-0.31*** (0.11)	-0.17 (0.17)
Class size	0.24 (0.58)	0.03 (0.58)

\*  $p < 0.10$ , \*\*  $p < 0.05$ , \*\*\*  $p < 0.01$

The table reports the coefficient from regressions of each outcome variable on standardized class ability (excluding student  $i$ ). Column 1 corresponds to equation 3.3, column 2 adds school FE and corresponds to equation 3.4. Standard errors are clustered at the class level.



Table 3.3.2:  
Does peer ability predict parental occupation?

Outcome	OLS	School FE
non-skilled workers, goods	-0.04*** (0.01)	0.02 (0.01)
non-skilled workers, service	-0.04*** (0.01)	0.00 (0.02)
skilled workers, goods	-0.06*** (0.01)	-0.03 (0.02)
skilled workers, service	-0.01 (0.00)	-0.01 (0.01)
lower non-manual ii	-0.01 (0.01)	-0.01 (0.01)
lower non-manual i	-0.01* (0.01)	-0.01 (0.01)
intermediate-level non-manual	0.05*** (0.02)	-0.01 (0.02)
higher civil servants and senior salaried	0.12*** (0.02)	0.04*** (0.01)
entrepreneur	0.02** (0.01)	0.01 (0.02)
farmer	0.01 (0.01)	0.02 (0.01)

\*  $p < 0.10$ , \*\*  $p < 0.05$ , \*\*\*  $p < 0.01$

The table reports the coefficient from regressions of each outcome variable on standardized class ability (excluding student *i*). Column 1 corresponds to equation 3.3, column 2 adds school FE and corresponds to equation 3.4. Standard errors are clustered at the class level.

### 3.3.3 Results

Table 3.3.3 summarizes the main empirical results. I report the coefficient  $\beta$  from equation 3.2 for each outcome, under different specifications. Specification (1) does not include controls apart from the pool correction (here school peer ability) suggested in Guryan et al. (2009). Specification (2) adds a control for own ability. Specification (3) is my preferred specification, and adds individual level controls for gender, parental occupation (10 dummies), foreign parents and special education. Specification (4) adds class level controls: number of teacher changes, class size, and fraction of students without ability data in the class. This specification is added as a robustness check. Class level controls might correlate with class ability, and are not necessarily exogenous. While they might introduce some bias into the coefficient of interest, it is still important to check whether the coefficient of interest changes significantly after controlling for these variables. Finally the last column reports the mean of each outcome variable for the estimation sample as a reference.

Before interpreting results it is important to highlight again some points. First, the fixed effects specification uses the subsample of schools where at least two classes were sampled. In section 3.2.3 I showed that this restricted sample is very similar to the final sample. Sample selection should not be an issue. However I use half of the original sample, and exploit only within school variation in average class ability. This negatively affects the precision of the coefficient of interest, which implies I cannot detect small effects. Second, I normalize all ability measures at the cohort level. When interpreting the effect of higher peer ability, it is important to recall that the standard deviation of mean class ability is around 40% of the national standard deviation (see Table 3.2.1). Lastly, I cluster standard errors at the level at which the “treatment” changes, that is at the class level.

I find a statistically significant negative effect of peer ability on the choice of advanced math courses. This effect appear to persist over late compulsory school, and is stable across specifications. An increase of one standard deviation in class ability reduces by 5.6 percentage points the probability that a student takes an advanced course in math. This is a nontrivial effect, and corresponds to a 2.2 percentage points reduction for a one standard deviation increase of average class ability. Class ability instead does not significantly affect the choice of an advanced course in English. The point estimate is negative, as expected, but much smaller than the coefficient for math. Finally, peer ability does not

seem to affect high school track choice.<sup>11</sup>

Table 3.3.3: Main results

Outcome variable	Class ability coefficient				Mean
	(1)	(2)	(3)	(4)	
Advanced math in grade 7	-0.065** (0.027)	-0.069*** (0.027)	-0.056** (0.027)	-0.056** (0.027)	0.752
Advanced math in grade 8	-0.068*** (0.024)	-0.073*** (0.023)	-0.064*** (0.023)	-0.064*** (0.023)	0.665
Advanced math in grade 9	-0.060** (0.030)	-0.066** (0.029)	-0.060** (0.029)	-0.056** (0.028)	0.563
Advanced English in grade 7	-0.027 (0.026)	-0.031 (0.026)	-0.017 (0.024)	-0.017 (0.024)	0.765
Advanced English in grade 8	-0.014 (0.023)	-0.018 (0.022)	-0.005 (0.021)	-0.008 (0.021)	0.735
Advanced English in grade 9	-0.036 (0.022)	-0.041* (0.022)	-0.026 (0.022)	-0.030 (0.021)	0.686
Academic track in high school	-0.008 (0.035)	-0.012 (0.034)	-0.014 (0.028)	-0.017 (0.027)	0.554

\*  $p < 0.10$ , \*\*  $p < 0.05$ , \*\*\*  $p < 0.01$

The table reports coefficients from a regression of each outcome variable on standardized class ability (excluding student  $i$ ), controlling for school fixed effects. Specification (1) does not include any controls apart from the pool correction (here school peer ability) suggested in Guryan et al. (2009). Specification (2) controls for own ability. Specification (3) adds individual level controls (parental occupation dummies, foreign parents, gender, special education status) to specification (2). Specification (4) adds class level controls (teacher changes, class size, fraction of the class without ability data) to specification (3). Standard errors are clustered at the class level.

<sup>11</sup>Facchinello (2016) shows that high performance and academic choices in late compulsory school do not necessarily translate into choosing academic tracks in high school. Controlling for ability, enrollment into academic tracks appears to be strongly related to SES.

### 3.4 Mechanisms

In this section I present additional results, in order to better understand which of the potential mechanisms outlined in Section 3.3.1 explain the main findings. I summarize these results in Table 3.4.1, which follows the same structure of Table 3.3.3. All outcomes, apart from the test scores, are derived from the survey questions given by students in grade 6.<sup>12</sup>

#### 3.4.1 Beliefs

The first two rows of Table 3.4.1 relate in different ways to student self-assessment. *Evaluation of skills in class* is the answer to a question asking students to rank themselves in their class on a 1-9 scale. Students are asked to assess their skills, rather than their ability in class, so I do not expect a one to one relationship with average class ability. Finding an effect on this variable simply implies that students are aware of their peers' performance. This is a necessary condition for any mechanism that involves students assessing themselves with reference to their peers. I confirm that average peer ability negatively affects how students rank their school performance *within the class*. The effect is stable over all the specifications. An increase of a standard deviation in ability reduces student evaluation by half a point in a 1 to 9 scale. This corresponds to one fifth of a point reduction for an increase of one standard deviation in average class ability (about half of the between class standard deviation of the outcome).

The previous result does not imply that students' beliefs about own ability are affected by peer's ability. I can test whether this is the case looking at the effect of peer ability on the variable *Evaluation of skills*, an index based on a battery of yes/no questions asking students whether they perform well in different domains (sums, spelling, reading, school in general). I find a negative and statistically significant effect, comparable in magnitude to the one estimated for *Evaluation of skills in class*. This confirms that students assess their skills in relation to their peers.

---

<sup>12</sup>See Section 3.2.1 for a description, and Appendix 3.A for more details on the indexes used.

### 3.4.2 Performance feedbacks

Another important channel through which peer ability might affect choices is academic performance. Positive peer effects in school performance could positively affect the choices of students. This effect works thus in the opposite way of the belief channel. It is important to assess this effect in the light of the different results I find in math and English choices: do different peer effects in performance explain the different effects I find? Rows 3 and 4 of Table 3.4.1 suggest this might be the case. I find statistically significant positive effects in English, but not in Math performance. The magnitude of the coefficient in English is much larger than the one in math: a one standard deviation increase in class average ability leads to a 7% standard deviation increase in the national English test score. This suggests that the two subjects are qualitatively different: there are more positive spillovers in performance in English rather than math. It might simply be easier to learn from better peers when studying a language - which generally involves group work and conversation - rather than when studying math, where learning should be more individual.

### 3.4.3 Motivation and imitation

Peer ability might also affect students through changes in student preferences and willingness to put effort in school. This in turn relates to two conceptually different mechanisms. On the one hand, having better students in class might trigger imitation, and thus positively affect school engagement. On the other hand if students' beliefs about skills are positively related to school engagement, the effect might go in the opposite direction. I look at the overall effect by building a motivation index, scaled 0-9, using yes/no survey questions asking students how engaged they were in school. Row 5 of Table 3.4.1 suggests that peer ability does not affect motivation. Notice however that all the regressions from which this effect was estimated exhibit very low  $R^2$ , and thus also suffer more from precision issues.

### 3.4.4 Classmates, teacher and parental responses

Finally I look at whether students report in the surveys different behavior from their classmates, parents, and teachers, when assigned to higher ability classes. It is hard to establish the direction in which peer ability should affect the behavior of these important inputs (see the discussion in section 3.3.1), so these empirical

findings are quite interesting to better understand mechanisms.

I find that higher ability classes do not significantly affect students' interactions with their classmates. They also do not seem to affect parental support. I find instead that higher peer ability positively affects teachers' helpfulness. A class with smarter children might be easier to manage and to teach. This could improve teacher effectiveness for all students, independently of their ability. Notice that the effect could have gone in the opposite direction if teachers adapted to the average ability of the class. In the setup considered this is not very likely: in this period national curricula establish in detail the educational goals the class has to achieve.

Table 3.4.1: Mechanisms

Outcome variable	Class ability coefficient				Mean
	(1)	(2)	(3)	(4)	
Evaluation of skills in class (1-9)	-0.524*** (0.078)	-0.539*** (0.077)	-0.522*** (0.078)	-0.522*** (0.078)	5.770
Evaluation of skills (1-9)	-0.598*** (0.163)	-0.610*** (0.162)	-0.557*** (0.160)	-0.554*** (0.162)	6.711
Math std test score, Grade 9	0.105 (0.082)	0.099 (0.079)	0.110 (0.078)	0.083 (0.080)	0.013
English std test score, Grade 8	0.178** (0.074)	0.176** (0.074)	0.184** (0.073)	0.151** (0.073)	0.014
Motivation in school (0-9)	-0.132 (0.134)	-0.141 (0.135)	-0.131 (0.134)	-0.104 (0.131)	5.678
Class environment (1-9)	0.064 (0.101)	0.061 (0.100)	0.081 (0.098)	0.095 (0.096)	7.935
Teacher helpfulness (0-9)	0.309** (0.152)	0.301** (0.148)	0.316** (0.145)	0.343** (0.148)	7.031
Help at home with homework	0.024 (0.025)	0.025 (0.025)	0.029 (0.025)	0.034 (0.024)	0.801

\*  $p < 0.10$ , \*\*  $p < 0.05$ , \*\*\*  $p < 0.01$

The table reports coefficients from a regression of each outcome variable on standardized class ability (excluding student  $i$ ), controlling for school fixed effects. Specification (1) does not include any controls apart from the pool correction (here school peer ability) suggested in Guryan et al. (2009). Specification (2) controls for own ability. Specification (3) adds individual level controls (parental occupation dummies, foreign parents, gender, special education status) to specification (2). Specification (4) adds class level controls (teacher changes, class size, fraction of the class without ability data) to specification (3). Standard errors are clustered at the class level. *Evaluation of skills in class* is the answer to a question asking where the student is in terms of skills in the class distribution. *Evaluation of skills*, *Motivation*, *Class environment* and *Teacher helpfulness* are indexes built from survey questions asking respectively students: to evaluate own ability in different domains (math, spelling, reading, etc), their motivation in school (if they put effort in school tasks, if they give up easily, etc), availability and helpfulness of the teacher, and how they get along with their classmates.

### 3.5 Conclusion

In this paper I use detailed survey data on a cohort of Swedish 6th graders to estimate the effect of classmates' ability on students' choices in compulsory school. I find that a one standard deviation increase in average class ability reduces by 2 percentage points the probability of taking an advanced math course in grades 7 to 9. Peer ability does not affect English course choices in grades 7 to 9, and whether students choose academic tracks in high school.

I look at underlying mechanisms and show evidence that peer ability negatively affects students' assessment of own ability. This effect is known in the psychology literature as the big-fish-in-a-little-pond effect (Marsh, 1987). The different effects I find on math and English course choice appear to reflect different spillovers in performance: I show that students benefit much more from having high ability peers in English, an interactive subject, than they do in math. Finally peer ability does not seem to affect student's motivation, class interaction and parental support, but positively affects teacher interaction.

The core literature on peer effects in education, recently reviewed by Sacerdote (2011), finds either positive, zero or non-linear effects of peer performance on individual school achievement. Consistently with this literature I find positive peer effects on academic performance only for one subject.

Recent work has focused on a different channel, and shows that school rank, independently of peer ability, can negatively affect student performance and academic choices.<sup>13</sup> My results are consistent with Elsner & Isphording (2015), who find that ranking higher in high school significantly affects perceived intelligence and teacher support, but I do not find a positive effect on long-run academic choices. I find no evidence of an overall increase in student confidence ("motivation" in my analysis), as in Murphy & Weinhardt (2014), or lower academic performance, as predicted by Tincani's model of rank concerns. I cannot however exclude that these two effects might be at play, when keeping ability fixed. All the cited papers indeed look at a pure rank effect. If peer ability affects performance in the opposite direction, as shown in my results, the overall effect on performance and motivation can systematically differ.

The contribution of the paper is indeed estimating the *net* effect of peer ability on academic choices. Ultimately this is the relevant parameter policy

---

<sup>13</sup>See Murphy & Weinhardt (2014), Tincani (2015), and Elsner & Isphording (2015).



makers should consider when designing tracking and streaming policies.

### 3.A Indexes

This section presents in detail the indexes I use as outcome variables in the analysis.

*Evaluation of skills in class* is the answer to a survey question asking students to evaluate on a 1 to 9 ordinal scale their position in the class skill distribution. If students evaluated themselves in an unbiased way the distribution would be normal. There is instead a tendency to overrate own skills within the class.

*Evaluation of skills* is a 1-9 index built from 4 yes/no survey questions, each one assigned 2 points, with the scale properly reversed when the question implies low skills: do you think that you are good at sums? (+); do you think that you are good at spelling? (+); do you think that you are bad at reading? (-); do you think that you do well in school? (+). The distribution is concentrated on high values, but many students still answer positively only to half of the questions.

*Motivation* is a 0-9 index built from 6 yes/no survey questions, each one assigned 1.5 points, with the scale properly reversed when the question implies low motivation: do you give up if you get a difficult task to do in school? (-); do you often think about other things when you do maths and writing in school? (-); do you think that it is unpleasant to have to answer questions in school? (-); do you find it difficult to give the right answer, even if you know it? (-); do you always do your best, even when the tasks are boring? (+); do you get disappointed if you get bad results in a test? (-). While the majority of students answer in a positive way to at least 3 items out of 6, there is a nontrivial fraction of disengaged/demotivated students.

*Class Environment* is a 1-9 index built from 4 yes/no survey questions, each one assigned 2 points, with the scale properly reversed when the question implies low interaction with other classmates: do you like working together with other children in the class? (+); do you think that other children in the class like working together with you? (+); do you worry about things that happen in school? (-); do you often spend time on your own during breaks? (-). The distribution is clearly concentrated on high values, with few students answering positively only to some questions.

*Teacher helpfulness* is a 0-9 index built from 6 yes/no survey questions, each

one assigned 1.5 points, with the scale properly reversed when the question reflects low teacher helpfulness/availability: do you ask the teacher for help when you do not understand? (+); would you like to ask the teacher for help more often than you do? (-); do you think that your teacher cares about you? (+); do you often think that the teacher should care more about you? (-); do you think that it is hard to understand when the teacher explains things for all the class (-); Do you often think that you would like to understand things better when the teacher explains things? (-). While the majority of students answer in a positive way to at least 4 items out of 6, there is a nontrivial fraction of students who answers positively to only three or less answers.

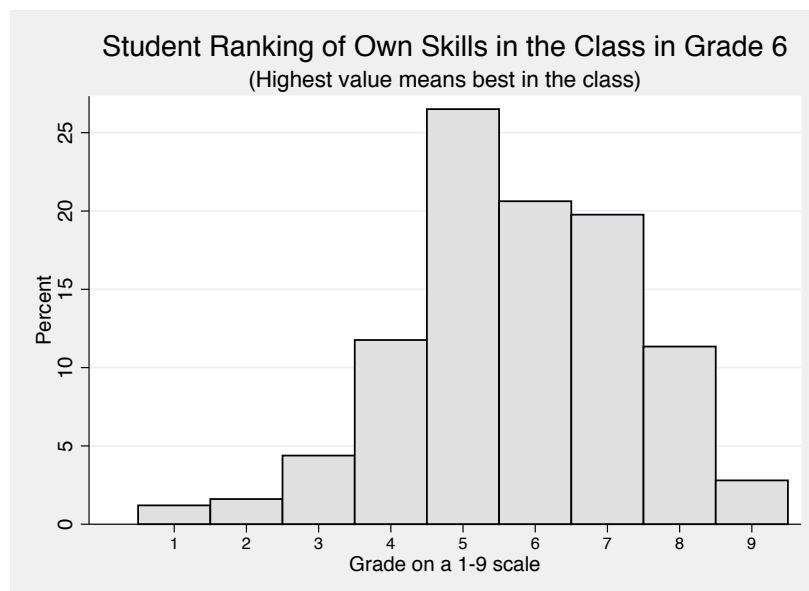


Figure 3.A.1: Distribution of *Evaluation of skills in class*

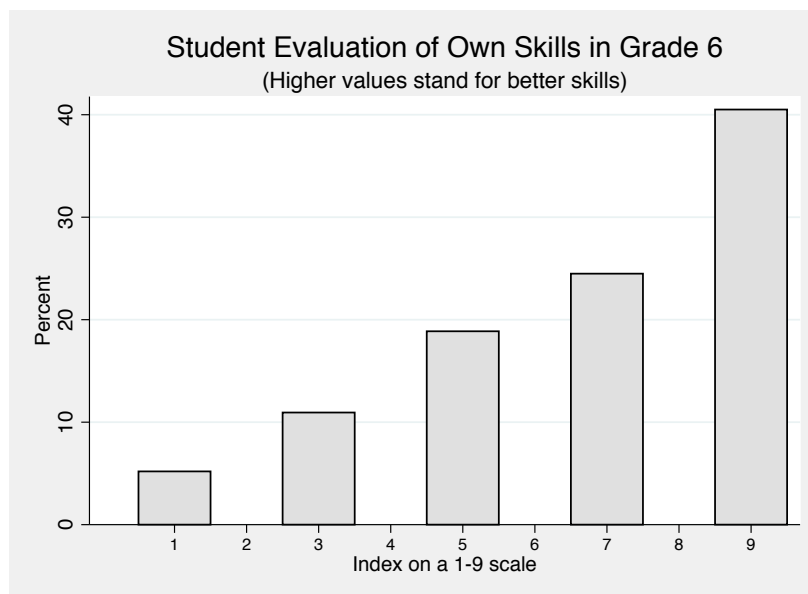


Figure 3.A.2: Distribution of *Evaluation of skills* index

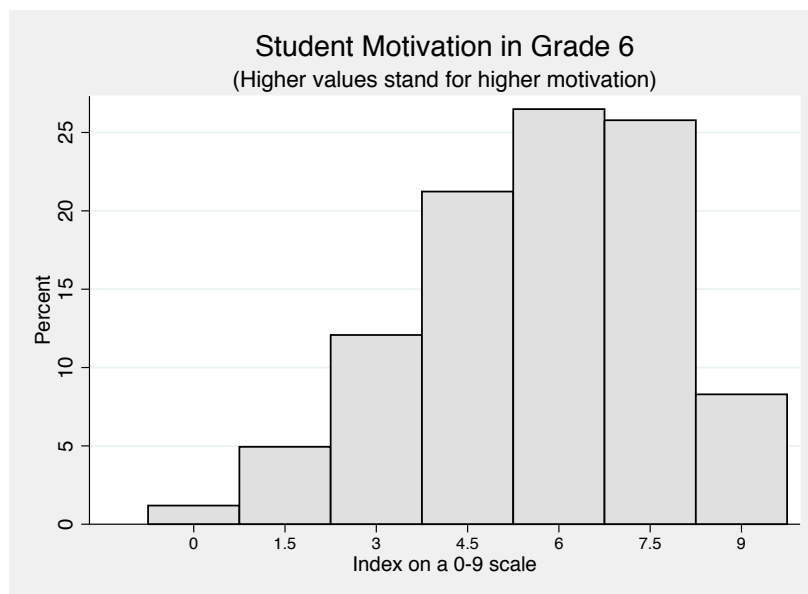


Figure 3.A.3: Distribution of *Motivation* index

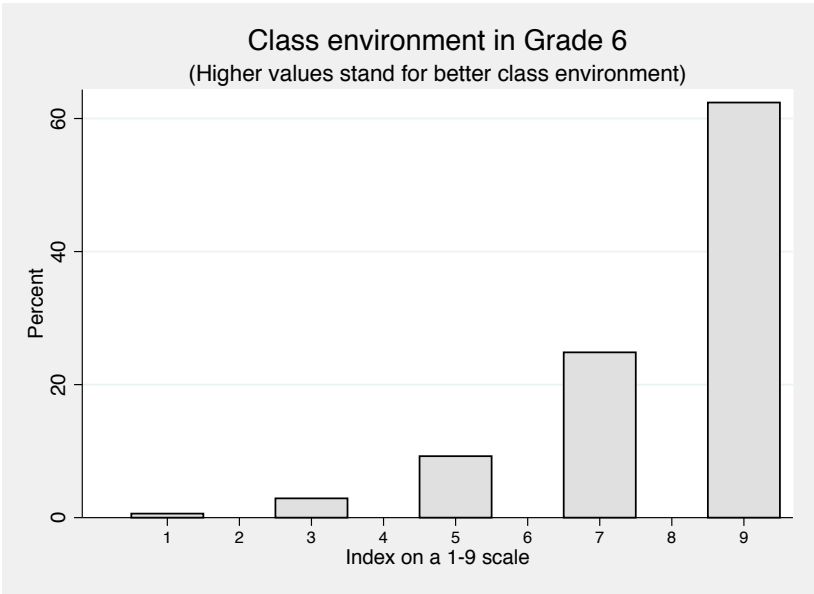


Figure 3.A.4: Distribution of *Class Environment* index

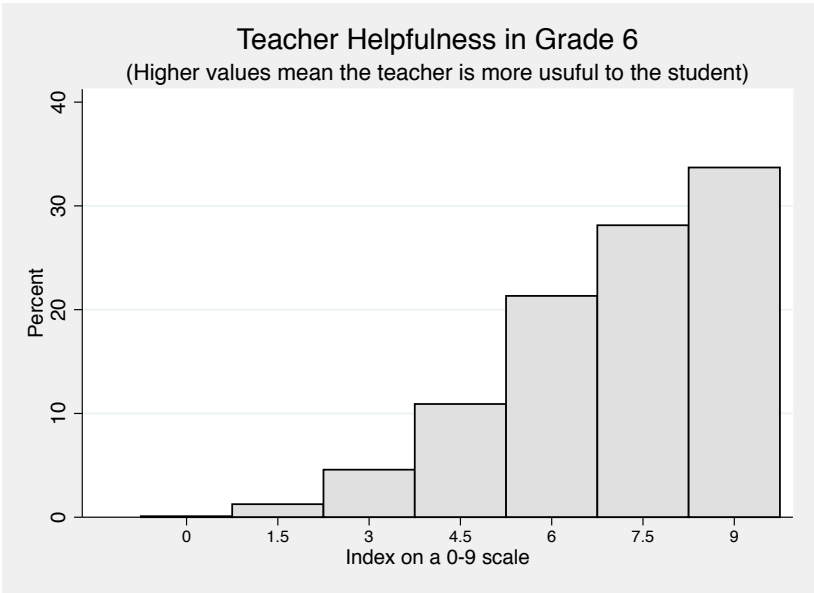


Figure 3.A.5: Distribution of *Teacher helpfulness* index

### 3.B Rank Deviations

Figure 3.B.1 shows that a substantial fraction of students exhibits nontrivial deviations between their percentile rank within the class and in the national distribution. If students form their beliefs about own ability with reference to their class, their evaluations might be thus biased. Notice that this measure is to some extent affected by class size, as in smaller classes percentile rank will mechanically differ from national rank, due to the different support of the variable.

While the previous picture shows deviations in ranking with respect to the national distribution, it is important to understand whether class ranking would differ if students were allocated to different classes in the same school. One way to see that is to consider the students in the restricted sample, and measure their ability ranking when randomly assigned to a different class in the same school. Figure 3.B.2 plots the difference between actual and simulated ranking for these students. There appears to be still a significant fraction of students whose ranking nontrivially changes due to different class ability composition. Notice that in this case the problem of class size is strongly reduced, due to the fact that there is less variation in class size within school (the average within-school standard deviation in class size is 1.8).

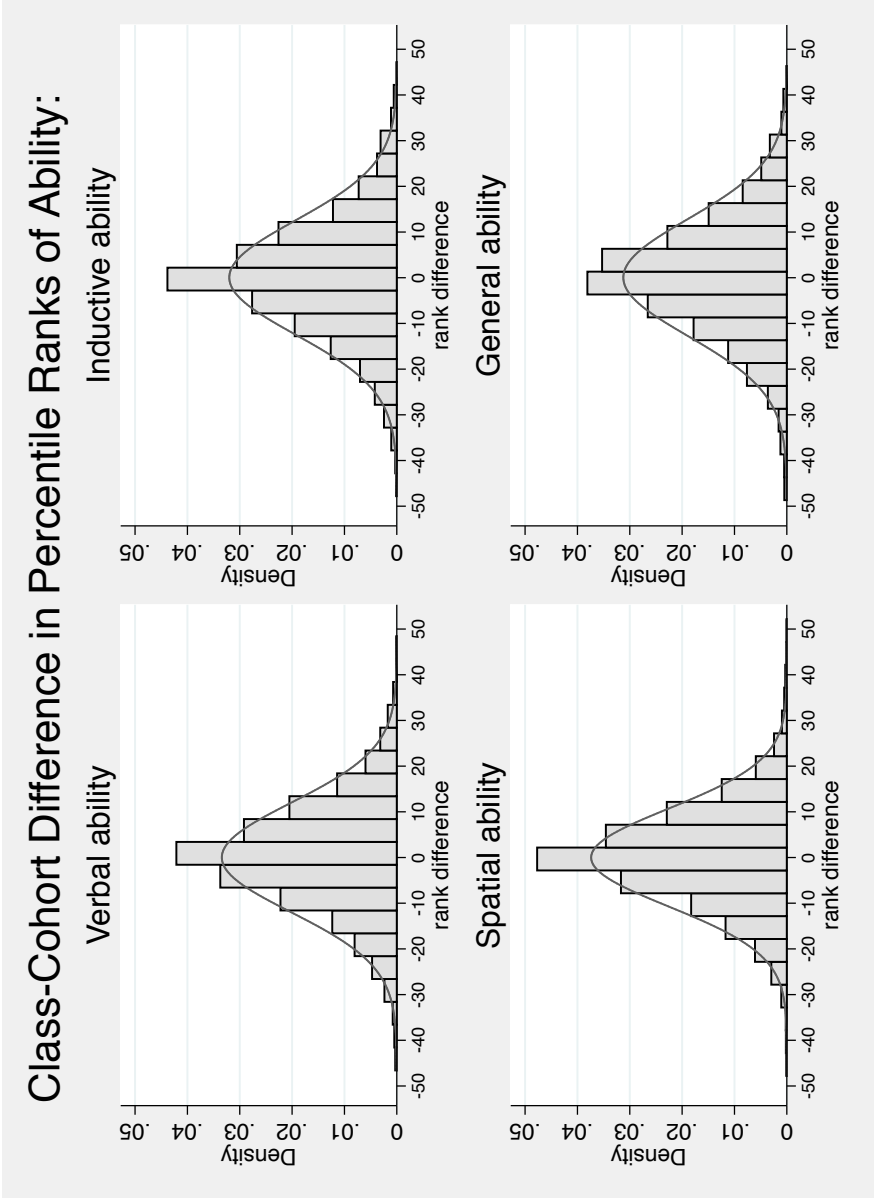


Figure 3.B.1: Comparison of class and national ability ranks

Note: The Figure plots the difference between the rank of the student in inductive, verbal, spatial and general ability at the class and at the cohort level. This is showing that students who rank high in their classes might rank much lower nationally, and the opposite. If students form their beliefs about own ability with reference to their class, their evaluations might be thus biased.

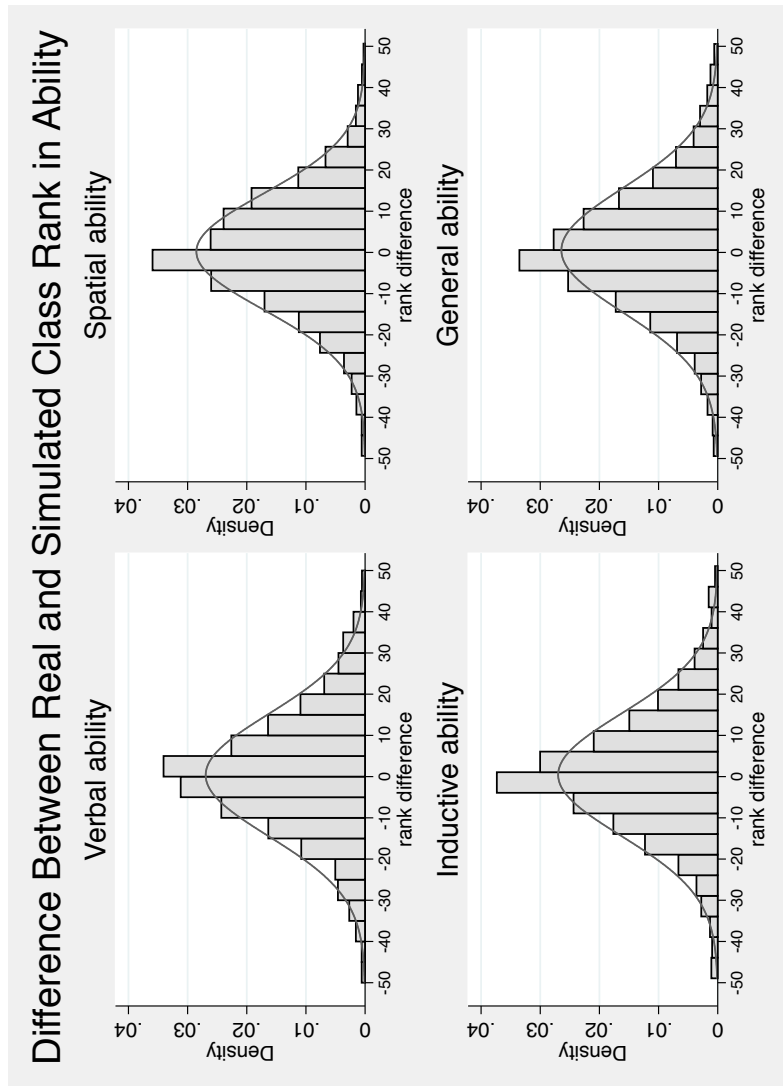


Figure 3.B.2: Comparison of actual class ability rank and rank when the student is moved to a different class in the same school

Note: The Figure plots the difference between the actual class rank of students in inductive, spatial, verbal and general ability, and the rank students would have if they were moved to a random class in the same school. This is showing that students who rank high in their class level might rank much lower in a different class, and the opposite. If students form their beliefs about own ability with reference to their class, their evaluations might be thus biased.

## References

- Abdulkadiroglu Atila & Angrist D. Joshua & Pathak A. Parag, 2014. "The Elite Illusion: Achievement Effects at Boston and New York Exam Schools," *Econometrica*, 82(1), 137-196
- Ammermueller Andreas & Pischke Jörn-Steffen, 2009. "Peer Effects in European Primary Schools: Evidence from the Progress in International Reading Literacy Study," *Journal of Labor Economics*, University of Chicago Press, vol. 27(3), pages 315-348, 07.
- Bobba Matteo & Frisanchi Veronica, 2014. "Learning About Oneself: The Effects of Signaling Academic Ability on School Choice". Inter-American Development Bank, mimeo.
- Borghans Lex & Meijers Huub & Weel Bas Ter, 2008. "The Role Of Noncognitive Skills In Explaining Cognitive Test Scores," *Economic Inquiry*, Western Economic Association International, vol. 46(1), pages 2-12, 01.
- Burke A. Mary & Sass A. Tim, 2013. "Classroom Peer Effects and Student Achievement," *Journal of Labor Economics*, 31(1), 51-82.
- Cunha Flavio & Heckman James, 2009. "The Economics and Psychology of Inequality and Human Development," *Journal of the European Economic Association*, MIT Press, vol. 7(2-3), pages 320-364, 04-05.
- Cunha Flavio & Heckman James, 2007. "The Technology of Skill Formation," *American Economic Review*, American Economic Association, vol. 97(2), pages 31-47, May.
- Elsner Benjamin & Ingo E., 2015. "A Big Fish in a Small Pond: Ability Rank and Human Capital Investment," *IZA Discussion Papers* 9121, Institute for the Study of Labor (IZA).
- Emanuelson Ingemar, 1979. "Utvärdering Genom Uppföljning av Elever - Ett Nytt Individual-statistikprojekt," Stockholm School of Teacher Education, Report 11-1979



Facchinello Luca, 2016. "The Impact of Early Grading on Academic Choices: Mechanisms and Social Implications," working paper, Stockholm School of Economics.

Fruehwirth Jane Cooley, 2014. "Can Achievement Peer Effect Estimates Inform Policy? A View from Inside the Black Box," *The Review of Economics and Statistics*, MIT Press, vol. 96(3), pages 514-523, July.

Guryan Jonathan & Kroft Kory & Notowidigdo J. Matthew, 2009. "Peer Effects in the Workplace: Evidence from Random Groupings in Professional Golf Tournaments," *American Economic Journal: Applied Economics*, American Economic Association, vol. 1(4), pages 34-68, October.

Kling R. Jeffrey & Liebman B. Jeffrey & Katz F. Lawrence, 2007. "Experimental Analysis of Neighborhood Effects," *Econometrica*, Econometric Society, vol. 75(1), pages 83-119, 01.

Markman M. Jacob & Hanushek A. Eric & Kain F. John & Rivkin G. Steven, 2003. "Does Peer Ability Affect Student Achievement?," *Journal of Applied Econometrics*, John Wiley & Sons, Ltd., vol. 18(5), pages 527-544.

Härnqvist Kjell, 1998. "A Longitudinal Program for Studying Education and Career Development," Report 1998: 01, Göteborgs universitet, Institutionen för pedagogik (1998).

Imberman Scott & Kugler D. Adriana & Sacerdote Bruce, 2012. "Katrina's Children: Evidence on the Structure of Peer Effects from Hurricane Evacuees," *American Economic Review*, 102(5), 2048-2082.

Jalava Nina & Joensen Schrøter Juanna & Pellas Elin, 2015. "Grades and rank: Impacts of Non-financial Incentives on Test Performance," *Journal of Economic Behavior & Organization*, Elsevier, vol. 115(C), pages 161-196.

Lavy Victor & Schlosser Analia, 2011. "Mechanisms and Impacts of Gender Peer Effects at School," *American Economic Journal: Applied Economics*, pp. 1-33.

Lavy Victor & Silva Olmo & Weinhardt Felix, 2012a. "The Good, the Bad and the Average: Evidence on Ability Peer Effects in Schools," *Journal of Labor Economics*, 30(2), 367-414.

Lavy Victor & Paserman M. Daniele & Schlosser Analia, 2012b. "Inside the Black Box of Ability Peer Effects: Evidence from Variation in the Proportion of Low Achievers in the Classroom," *The Economic Journal*, 122(559), 208–237.

Marsh W. Herbert, 1987. "The Big-fish-little-pond Effect on Academic Self-concept," *Journal of Educational Psychology*, 79(3), 280.

Murphy Richard & Weinhardt Felix, 2014. "Top of the Class: The Importance of Ordinal Rank," *CESifo Working Paper Series* 4815, CESifo Group Munich.

Öckert Björn, 2002. "Do University Enrollment Constraints Affect Education and Earnings?," *Working Paper Series* 2002:16, IFAU - Institute for Evaluation of Labour Market and Education Policy.

Sacerdote Bruce, 2011. "Peer Effects in Education: How Might They Work, How Big are They and How much do we Know Thus Far?," *Handbook of the Economics of Education*, 3, 249-277.

Sjögren Anna, 2010. "Graded children – Evidence of Longrun Consequences of School Grades from a Nationwide Reform," *Working Paper Series* 2010:7, IFAU - Institute for Evaluation of Labour Market and Education Policy.

Stinebrickner Todd & Stinebrickner Ralph, 2012. "Learning about Academic Ability and the College Dropout Decision," *Journal of Labor Economics*, University of Chicago Press, vol. 30(4), pages 707 - 748.

Stinebrickner Todd & Stinebrickner Ralph, 2014. "A Major in Science? Initial Beliefs and Final Outcomes for College Major and Dropout," *Review of Economic Studies*, 81(1), 426 472.

Sund Krister, 2009. "Estimating Peer Effects in Swedish High School Using School, Teacher, and Student Fixed Effects," *Economics of Education Review*, Elsevier, vol. 28(3), pages 329-336, June.

Tincani M. Michela, 2015. "Heterogeneous Peer Effects and Rank Concerns: Theory and Evidence," working paper.

Zafar Basit, 2011. "How Do College Students Form Expectations?," *Journal of Labor Economics*, University of Chicago Press, vol. 29(2), pages 301 - 348.